

[illegible][illegible]


```

DR 8MBU, AC011713; AAF14654.1; -.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40. 4.
DR PRINTS: SM00320; GPROTEINRPT.
DR PROSITE: PS00678; WD40.
DR PROSITE: PS00892; WD_REPEATS_2; 3.
DR PROSITE: PS00294; WD_REPEATS_REGION; 1.
DR Hypothetical protein; Repeat; WD repeat.
DR SW 70 100 WD 5.
FT REPEAT 112 142 WD 3.
FT REPEAT 153 181 WD 4.
FT REPEAT 251 AA; 27677 MW; 39780AC60E240F CRC64;
SQ SEQUENCE 251 AA; 27677 MW; 39780AC60E240F CRC64;

Query Match
Best Local Similarity 19.98; Score 352.5; DB 1; Length 251;
Matches 86; Conservative 45; Mismatches 87; Indels 61; Gaps 6;

Cc Db 4 SNEFKLQPPEDISSVFNFTSQPLVSSNVTGLVQVP-----ANSRLKYQHTG 57
15 NGSYGVTFEAFSSLSGFSPFR-ADILWATSNMNVQVCHZISRGASLAGAPKASTISHQ 73
58 AVIDCAVDTP-THANGSGDQKQHWLD-LNTDQENLVGTHDAPICVCEPVNVMWTS 115
74 PVLCSANKDDDTVFSGDCQDAKAWPILSGSGQVTVNAGEGTFAMANTFOMKLLATS 133
116 WDDTVLWLDLTPFCNAGTFSPQEKVTVLSSGDLVITAGRVLLWLDLNNQVQGRE 175
134 NQKTLKLVTPQNFVTHQQLQDCKYLSVPLVPLVGVGTADRLVLFNPLQNTPEFKI 193
176 SSLKYQVTRCTAFRPNQVQY-VLSSIGSRVAVLVDPSVEQKKVAKFKHLKENNIO 233
194 SPLKLVQTRVTFAPDQSGFLKSLADAVIL----- 224
234 IYPNMAISFNHINTATGSGDQSPNPNKVLQIQP 272
225 -----DGAENFWDKSGQLANVF 242

RESULT 9
ID Y146 ANASP STANDARD; PRT; 1526 AA.
CD 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DE Hypothetical WD-repeat protein alr3466.
OS Anabaena sp. strain PCC 7120.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
CX NCBI_TaxId=103690;
RL MEDLINE=21595285; PubMed=11759840;
RX Kaneko T., Makamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Katsube A., Iriguchi M., Shikawa A., Kawashima K., Kimura T.,
RA Katsube A., Iriguchi M., Shikawa A., Kawashima K., Kimura T.,
RA Nakasaki M., Shimo S., Sugimoto M., Takawara N., Tamada N.,
RA Yasuda M., Tabata S.;
RL "Complete genomic sequence of the filamentous nitrogen-fixing
RT DNA bacterium JCM-21312 (2803) sp. strain PCC 7120.";
CC -/- SIMILARITY: Contains 16 WD repeats.
CC -/- SIMILARITY: Contains 1 pentapeptide repeat domain.
CC This entry was produced through a collaboration
CC between the Swiss Institute of Bioinformatics and
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC derived from the EMBL database. Usage by for-profit commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/embldb/)

```

```

or send an email to license@ebi.ac.uk.
Cc Db 8MBU, AC011713; AAF14654.1; -.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40. 4.
DR PRINTS: SM00320; GPROTEINRPT.
DR PROSITE: PS00678; WD40.
DR PROSITE: PS00892; WD_REPEATS_2; 3.
DR PROSITE: PS00294; WD_REPEATS_REGION; 1.
DR Hypothetical protein; Repeat; WD repeat.
DR SW 70 100 WD 5.
FT REPEAT 112 142 WD 3.
FT REPEAT 153 181 WD 4.
FT REPEAT 251 AA; 27677 MW; 39780AC60E240F CRC64;
SQ SEQUENCE 251 AA; 27677 MW; 39780AC60E240F CRC64;

Query Match
Best Local Similarity 12.94; Score 228; DB 1; Length 1526;
Matches 85; Conservative 57; Mismatches 121; Indels 66; Gaps 11;

Cc Db 17 ISSVSVTSQPLVSSNVTGLVQVP-----ANSRLKYQHTGVLQCF-YQTHNMSG 74
951 VRSVTSFP-SMLASGSSDQTVRLMDISSCLYIFQHTGTVSVFVNLGSLMGTS 1009
76 LDHQLQKHLATQDQ-ENLVGTHDAPICVCEPVNVMWTSQHTQWLP- 124
1010 GQTVRLMDISSQCFYFQHTGTVSVFVNLGSLMGTS 1069
125 -PRTCNQNTQSGEK- 148
1070 LQHTGTVSVFVNLGSLMGTSQDQTVRLMDISSCLYIFQHTGTVSVFVNLGSLMGTS 1129
149 LQHTGTVSVFVNLGSLMGTSQDQTVRLMDISSCLYIFQHTGTVSVFVNLGSLMGTS 1207
1130 TLANGSDQVRLMDISS-----KCLYTLQHTNWNVAVSPDQTLASGSDQTVL 1184
208 LDPSVQKKVAKFKHLKENNIDQYVNVNMFHINTATGSGDQSPNPNK 267
1185 WDIS-----SSKLYLSDQTLASGSDQTVLMDISSK 1233
268 RLQHPYFYSIALAFSDQTLASG 296
1234 CLCTFQHTGTVSVFVNLGSLMGTS 1262

RESULT 10
ID Y500 ANASP STANDARD; PRT; 1258 AA.
CD 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DE Hypothetical WD-repeat protein alr2800.
OS Anabaena sp. strain PCC 7120.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
CX NCBI_TaxId=103690;
RL MEDLINE=21595285; PubMed=11759840;
RX Kaneko T., Makamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Katsube A., Iriguchi M., Shikawa A., Kawashima K., Kimura T.,
RA Katsube A., Iriguchi M., Shikawa A., Kawashima K., Kimura T.,
RA Nakasaki M., Shimo S., Sugimoto M., Takawara N., Tamada N.,
RA Yasuda M., Tabata S.;
RL "Complete genomic sequence of the filamentous nitrogen-fixing
RT DNA bacterium JCM-21312 (2803) sp. strain PCC 7120.";
CC -/- SIMILARITY: Contains 16 WD repeats.
CC -/- SIMILARITY: Contains 1 pentapeptide repeat domain.
CC This entry was produced through a collaboration
CC between the Swiss Institute of Bioinformatics and
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC derived from the EMBL database. Usage by for-profit commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/embldb/)

```

[illegible]

[illegible]

	PRT)	486	AA.
AD	PR1L ARATH	STANDARD;	
AC	Q42384;		
DT	15-JUN-1998	(Rel. 36,	Created)
DI	15-JUN-1998	(Rel. 36,	Last sequence update)
DD	18-SEP-2003	(Rel. 42,	Last annotation update)
DE	PP1/P2A phosphatases pleiotropic regulator PR1L		
DN	PR1 OR ARI51900 OR MLI590K		
GN	Arabidopsis thaliana (Cruciferae) crassae		
GO	Sukromatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophytia; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
CC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.		
NC	!_taxid=3702;		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=cv. Columbia;		
RF	Schell S., Funkhøj P., Stankovic B., Bakso L., Machur J., Redei G.P.,		
RH	Submitted (NOV-1994) to the EMBL/GenBank/DBS databases.		
RJ	[2]		
RD	SEQUENCE FROM N.A.		
RE	MEDLINE=9812317; PubMed=9461235;		
EX	Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,		
RA	Brigand R., Dirke W., van Staveren M., Stiekema W., Drost L.,		
RA	Kedler H., Wedler E., Wandorf R., Weitzbauer P., Hecht T., Torry N.,		
GA	Glien J., Villarroel R., De Clecq R., van Montagu M., Lecharny A.,		
RA	Aubourg S., Gy J., Kreis M., Lao N., Kavanagh T., Hempel S.,		
RA	Müller-Auerbach D., Rieger M., Schaefer M., Fink B.,		
RA	Puidgumench P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,		
RA	Cavandi E., Chemsater B., Hilbert H., Uesterhoefst A., Moores T.,		
RA	Corado V., Serey Delany F., Benes V., Reumann S., Anorge W.,		
RA	Closterman S., Schueller C., Chalvatzis N., Volckaert G., Meves H.,		
RT	"Analysis of 1.9 kb of contiguous sequence from chromosome 4 of		
RT	Arabidopsis thaliana";		
RI	Nature 351:483-488(1998).		
RI	[3]		
RD	SEQUENCE FROM N.A.		
RE	STRAIN=cv. Columbia;		
RF	MEDLINE=9810393; PubMed=10647199;		
RA	Mayer K.F.X., Schueller C., Wandorf R., Murphy G., Volckaert G.,		
RA	Pol T., Uesterhoefst A., Stiekema W., Entian K.-D., Terry N.G.,		
RA	Harris B., Anorge W., Brandt P., Grivell L.A., Rieger M.,		
RA	Kreis M., Milioni D., de Simone V., Oberbauer B., Mahe R., Mueller M.,		
RA	Reichart B., Portetleite D., Perez-Alonso M., Botry M., Bancroft I.,		
RA	Vos P., Hohsieil J., Zimmermann W., Wedler H., Ridley P.,		
RA	van der Schooten J., Gough B., Billham C., Kobben J., Vandebussche F.,		
RA	Braeken M., Weljens I., Voet M., Bastiaens I., Aert R., Defoor E.,		
RA	Maitzenegger T., Bothe G., Rampeger U., Hilbert H., Braun M.,		
RA	Wolter S., Brandt A., Peeters S., van Staveren M., Dirke W.,		
RA	Beinseiser S., Hempel S., Feldpausch M., Lambert S., Van den Daele H.,		
RA	De Keyser A., Buyshaert C., Glien J., Villarroel R., De Clecq R.,		
RA	van Montagu M., Rogers J., Cronin A., Quail M., Bray-Alley S.,		
RA	Patet A., Dandekar A., Lin S., Kay M., Lennard N., McCary R.,		
RA	Borkova D., Blocker H., Schärfe M., Grimm M., Lochner T.-H.,		
RA	Dose S., de Haan M., Maarze A.C., Schaefer M., Mueller-Auer S.,		
RA	Nabali C., Fuchs W., Farrar B., Ganderath K., Dauner D., Herzi A.,		
RA	Masseton O., Quigley F., Claibaud G., Muendlein A., Fabher R.,		
RA	Schnabl S., Hiller R., Schmidt M., Leckony A., Augsburg S.,		
RA	Creider F., Cooke R., Berger C., Meyer S., Trol J., Torres A.,		
RA	Perez-Perez A., Furnelle B., Bont E., Johnson S., Francis P., Bleke C.,		
RA	Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bleke C.,		
RA	Frimhan D., Haase D., Lemcke K., Meves H.-W., Stocker S.,		
RA	Parnelli J., Hedha N., Gronwald M., Schürmann X., Habermann X.,		

Qy 182 TRCIRA---FNNQOVLSTECRVAVVYLDPSPEVQKKYAFKCHLKENNIBQIY--- 235
Db 427 KYCLFANNSYPLGCHHISGSEDRIDW-----SLOTQK-----IYVLEIGH 469
Qy 236 --PVNAISFNNHNTFATG--SDGFVIM 261
Db 469 TTPVLATOSHPTLNIASSGLEPNNVIRI 498

Search completed: November 12, 2003, 20:09:20
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: November 12, 2003, 20:06:37 ; Search time 41 Seconds
(without alignments)
2064,421 Million cell updates/sec

Title: US-10-084-700-2
Perfect score: 1772
Sequence: 1 MTGSNEFLKQPPEDGISV.....DGFIROVTDATKRSPECT 328
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 01

Maximum Match 100

Listing first 45 summaries

Database : SPTREMBL 23.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.molluscs.*
- 8: sp.mono.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.protist.*
- 12: sp.reptile.*
- 13: sp.invertebrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.yeast.*
- 17: sp.archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1651	93.2	324	13	Q9a2h xenopus lae
2	1651	93.2	324	13	Q9a2h xenopus lae
3	1050.5	59.3	326	5	Q9a2h xenopus lae
4	1050.5	58.4	326	5	Q9a2h xenopus lae
5	949	53.6	327	5	Q9a2h xenopus lae
6	949	53.6	327	5	Q9a2h xenopus lae
7	949	53.6	327	5	Q9a2h xenopus lae
8	822.5	46.4	323	5	Q9a2h xenopus lae
9	710	40.1	343	5	Q9a2h xenopus lae
10	691	39.0	250	10	Q9a2h xenopus lae
11	671	37.9	327	5	Q9a2h xenopus lae
12	566	32.7	327	5	Q9a2h xenopus lae
13	541.5	30.6	349	10	Q9a2h xenopus lae
14	541.5	30.6	349	10	Q9a2h xenopus lae
15	530.5	29.9	349	3	Q9a2h xenopus lae
16	523.5	29.3	349	5	Q9a2h xenopus lae

17 521.5 29.4 346 5 Q9a2E7
18 518.5 29.3 320 3 Q42860
19 517.5 29.3 348 11 Q8C570
20 508 29.2 348 11 Q8C570
21 492 27.8 208 5 Q8B970
22 457 25.8 314 10 Q8LA29
23 452.5 25.5 315 10 Q8C795
24 452.5 25.5 315 10 Q8C795
25 345 18.3 336 5 Q9A170
26 325 18.3 336 5 Q9A170
27 316 17.8 318 5 Q8SRM6
28 316 17.8 318 5 Q8SRM6
29 231 12.0 335 11 Q8B142
30 231 12.0 335 11 Q8B142
31 206 11.6 1227 16 Q8Z0R1
32 199.5 11.2 140 10 Q8LMO1
33 198 11.2 140 10 Q8LMO1
34 136 11.1 358 5 Q9Y311
35 194.5 11.0 1747 16 Q8Z020
36 192 10.8 356 11 Q8C709
37 188 10.6 329 16 Q8X123
38 188 10.6 329 16 Q8X123
39 187 10.6 1189 16 Q8Y109
40 187 10.6 1551 16 Q8Y109
41 185.5 10.5 297 3 Q9C1X0
42 185.5 10.5 297 3 Q9C1X0
43 185.5 10.5 407 4 Q8B870
44 184.5 10.4 415 4 Q8N136
45 183.5 10.4 415 4 Q8N176

ALIGNMENTS

RESULT 1

Q9a2h2 ID Q9a2h2 PRELIMINARY; PRT; 324 AA.
NC Q9a2h2; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
MD repeat protein Bub3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
ON NCBI Taxonomy; Xenopus.
RN [1]-TaxID=9335.
RP SEQUENCE FROM N.A.
RX MEDLINE=21154833; PubMed=12311448;
RA Lewellyn, A., Malle, J., Gross S.D., Tunquist B.J., Tateb P.E.,
RT "Bub3 is activated by the protein kinase p90rsk during Xenopus oocyte maturation".
RT ENSE: NCBI TaxID=9335.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PF00320; GPROTINRPT.
DR SMART: SM00330; WD40; 1.
DR PROSITE: PS0082; WD_REPEATS_2; 3.
DR PROSITE: PS0294; WD_REPEATS_REGION; 2.
SQ Repeat; WD repeat.
Query Match 93.24; Score 1651; DB 13; Length 324;
Best Local Similarity 93.24; Pred. No. 2.8e-139;
Matches 302; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

Q9

1 MTGSNEFLKQPPEDGISV.....DGFIROVTDATKRSPECT 328
|||||

[illegible]

Query Match	53.08	Score 939	DB 10	Length 340
Best-Local	54.11	Pred. No. 13e-55		
Matches	178	Conservative 54	Mismatches 91	Indels 6
Gaps	2			
QY	1	MTGNEKRLKLPFGSTGVSQVDSVPSQFLVSSVTSVLSVYDANVNLKLYQGVAVL	60	
DB	4	VPSAGKELNPSFSGSLNLRFS--NSHSLDVLSSHDKVRLVYDVTNLSIKBEFLHGGAVL	62	
QY	61	CAKVPYTHANGSGVAGVYKRVFVNGVDEIYGHUKKVRGVSVYAGGVITSGNDYTK	120	
DB	63	DCCTHPSGSGVAGVYKRVFVNGVDEIYGHUKKVRGVSVYAGGVITSGNDYTK	122	
QY	121	KWPRF-----TCMAATGTSQSGVYKRVSVSRLLVNGVAGVYKRVSVSRLLVNGVAGV	175	
DB	123	KWDPRKAGSGPSTQVGTGLQPEVYSVSLVGHVRLVATAGVYKRVSVSRLLVNGVAGV	182	
QY	176	SELKYVTCIARFPMQGVLSISGRVAVELDPSEVQKRYKACHRLKXENTYVY	235	
DB	178	SELKYVTCIARFPMQGVLSISGRVAVELDPSEVQKRYKACHRLKXENTYVY	242	
QY	236	PVALVLFHNVHTATSGSDQVNVFPMKVRICPHRYVTFVLSIASNSQDTLLAS	295	
DB	243	PVNSIAFPHITVGTATGCGCPVNIWQGNKRLQVSKYNTLSLSFSRQDLAVAS	302	
QY	296	SVYVTEVQTEHPEDQITPIQVTDARTYK	324	
DB	303	SVYVTEGSGEPEALFVSNVLEVKR	331	
RESULT 8				
QY	Q8MSM6	PRELIMINARY	PRT	323 AA.
AC	Q8MSM6	01-OCT-2002 (T=EMBLrel. 22, Created)		
DT	01-OCT-2002	(T=EMBLrel. 22, Last sequence update)		
DT	01-MAR-2003	(T=EMBLrel. 23, Last annotation update)		
DE	L233540p.			
GN	B033.081	COY501		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Rhyngolidae; Drosophilidae; Drosophila.			
SN	[1]-TaxId=7227;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Barkeley;			
RA	Chenue M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise E.,			
RA	George R., Gonzalez M., Guarin H., Kromoller B., Li P., Liao G.,			
RA	Miranda A., Mungall C.J., Nunoo J., Paclab J., Pakasga V., Park S.,			
RA	Seibiniker S., Sumanamwong S., Wen K., Yu C., Lewis S.E., Hahn G.M.,			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY118526; AW49895.1;			
DR	NCBI; F0501180; W0400.			
DR	InterPro; IP001180; WD40.			
DR	PRINTS; PF00400; WD40; 4.			
DR	PROSITE; PRO0320; GROTEININFR.			
DR	PROSITE; PS00678; WD REPEATS 1; 1.			
DR	PROSITE; PS00682; WD REPEATS 2; 1.			
DR	PROSITE; PS00294; WD REPEATS REGION; 1.			
DR	REPEATS MD Repeat.			
DR	SEQUENCE 323 AA; 37515 MW; SARC3AB89515487 CRC64;			
Query Match	46.41	Score 822.5	DB 5	Length 323;
Best-Local	60.74	Pred. No. 3e-45;		
Matches	130	Conservative 40	Mismatches 56	Indels 1
Gaps	1			
QY	6	EFLNKPDPEDGVSIVKPSNT-SQFLVSSVTSVLSVYDANVNLKLYQGVAVL	60	
DB	5	EFLNKPDPEDGVSIVKPSNT-SQFLVSSVTSVLSVYDANVNLKLYQGVAVL	62	
QY	61	CAKVPYTHANGSGVAGVYKRVFVNGVDEIYGHUKKVRGVSVYAGGVITSGNDYTK	120	
DB	63	DCCTHPSGSGVAGVYKRVFVNGVDEIYGHUKKVRGVSVYAGGVITSGNDYTK	122	
QY	121	KWPRF-----TCMAATGTSQSGVYKRVSVSRLLVNGVAGVYKRVSVSRLLVNGVAGV	175	
DB	123	KWDPRKAGSGPSTQVGTGLQPEVYSVSLVGHVRLVATAGVYKRVSVSRLLVNGVAGV	182	
QY	176	SELKYVTCIARFPMQGVLSISGRVAVELDPSEVQKRYKACHRLKXENTYVY	235	
DB	178	SELKYVTCIARFPMQGVLSISGRVAVELDPSEVQKRYKACHRLKXENTYVY	242	
QY	236	PVALVLFHNVHTATSGSDQVNVFPMKVRICPHRYVTFVLSIASNSQDTLLAS	295	
DB	243	PVNSIAFPHITVGTATGCGCPVNIWQGNKRLQVSKYNTLSLSFSRQDLAVAS	302	
QY	296	SVYVTEVQTEHPEDQITPIQVTDARTYK	324	
DB	303	SVYVTEGSGEPEALFVSNVLEVKR	331	
RESULT 9				
QY	Q8MSM6	PRELIMINARY	PRT	343 AA.
AC	Q8MSM6	01-NOV-1999 (T=EMBLrel. 12, Created)		
DT	01-NOV-1999	(T=EMBLrel. 12, Last sequence update)		
DT	01-MAR-2003	(T=EMBLrel. 23, Last annotation update)		
DE	Y5409A.6	protein.		
GN	Y5409A.6	protein.		
OC	Eukaryota; Metazoa; Neomata; Chromodorea; Rhabditida; Rhabditidae;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
SN	[1]-TaxId=6239;			
RP	SEQUENCE FROM N.A.			
RA	Smye R.;			
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL032448; CA21699.1;			
DR	NCBI; F0001680; WD40.			
DR	InterPro; IP001180; WD40.			
DR	PRINTS; PRO0320; GROTEININFR.			
DR	PROSITE; PS00678; WD REPEATS 1; 1.			
DR	PROSITE; PS00682; WD REPEATS 2; 1.			
DR	PROSITE; PS00294; WD REPEATS REGION; 1.			
DR	REPEATS MD Repeat.			
DR	SEQUENCE 343 AA; 38127 MW; 021A00398GBDSE CRC64;			
Query Match	40.11	Score 710	DB 5	Length 343;
Best-Local	42.61	Pred. No. 3.7e-55;		
Matches	141	Conservative 67	Mismatches 109	Indels 10
Gaps	6			
QY	5	NEFLNKPDPEDGVSIVKPSNT-SQFLVSSVTSVLSVYDANVNLKLYQGVAVL	60	
DB	14	NEFLNKPDPEDGVSIVKPSNT-SQFLVSSVTSVLSVYDANVNLKLYQGVAVL	73	
QY	61	CAKVPYTHANGSGVAGVYKRVFVNGVDEIYGHUKKVRGVSVYAGGVITSGNDYTK	120	
DB	74	CAKVPYTHANGSGVAGVYKRVFVNGVDEIYGHUKKVRGVSVYAGGVITSGNDYTK	122	
QY	121	KWPRF-----TCMAATGTSQSGVYKRVSVSRLLVNGVAGVYKRVSVSRLLVNGVAGV	175	
DB	123	KWDPRKAGSGPSTQVGTGLQPEVYSVSLVGHVRLVATAGVYKRVSVSRLLVNGVAGV	182	
QY	176	SELKYVTCIARFPMQGVLSISGRVAVELDPSEVQKRYKACHRLKXENTYVY	235	
DB	178	SELKYVTCIARFPMQGVLSISGRVAVELDPSEVQKRYKACHRLKXENTYVY	242	
QY	236	PVALVLFHNVHTATSGSDQVNVFPMKVRICPHRYVTFVLSIASNSQDTLLAS	295	
DB	243	PVNSIAFPHITVGTATGCGCPVNIWQGNKRLQVSKYNTLSLSFSRQDLAVAS	302	
QY	296	SVYVTEVQTEHPEDQITPIQVTDARTYK	324	
DB	303	SVYVTEGSGEPEALFVSNVLEVKR	331	

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:07:13 Search time 20 Seconds

1577.166 Million cell updates/sec

Title: US-10-084-700-2

Hit score: 1 MTGSNEFKLPPGDISGV.....DGIIFQVTDATKPSRCT 328
 Sequence:

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283108 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 104

Listing first 45 summaries

Database:

PR-76:

1: p1r1*

2: p1r2*

3: p1r3*

4: p1r4*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	939	53.0	340	2	mitotic checkpoint
2	571.0	32.4	345	2	hypothetical protein
3	538	30.4	345	2	hypothetical protein
4	518.5	29.3	320	2	probable mitotic c
5	511.5	29.2	322	2	RNA export protein
6	511.5	29.2	322	2	hypothetical protein
7	498.5	28.1	341	1	cell cycle arrest
8	356.2	19.9	251	1	probable RNA export
9	352.5	19.9	251	2	polyA export prote
10	228.1	13.9	1526	2	WD-40 repeat prote
11	228.1	13.9	1526	2	WD-40 repeat prote
12	213.5	12.1	934	2	WD-40 repeat prote
13	209.1	11.8	1258	2	WD-repeat protein
14	198.6	11.5	1270	2	WD-repeat protein
15	198.6	11.5	1270	2	WD-repeat protein
16	194.5	11.0	1747	2	WD-40 repeat prote
17	187.1	10.6	1189	2	WD-repeat protein
18	187.1	10.6	1351	2	WD-repeat protein
19	187.1	10.6	1351	2	WD-repeat protein
20	178.5	10.1	1189	2	WD-repeat protei
21	177.5	10.0	356	2	hypothetical prote
22	176.5	10.0	677	2	serine/threonine k
23	176.5	10.0	677	2	PRK transducin-li
24	175.5	9.9	486	2	hypothetical prote
25	173.9	9.8	501	2	hypothetical prote
26	171.5	9.7	961	2	GTP-binding regula
27	169.5	9.6	340	2	GTP-binding regula
28	169.5	9.6	340	2	GTP-binding regula
29	169.5	9.6	340	2	G protein beta 1

ALIGNMENTS

RESULT 1

MTGS3286
 Mitotic checkpoint protein [imported] - Arabidopsis thaliana
 CSpecies: Arabidopsis thaliana (mouse-ear cress)
 CAccession: MTGS3286 #sequence_revision 24-Oct-2000 #feat_change 24-Oct-2000
 RName: T. Kato, T.; Sato, S.; Nakamura, Y.; Asamizu, E.; Tabata, S.
 submitted to the EMBL Data Library, September 1999
 AAccession: T52386
 AStatus: Preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Molecule length: 340
 A.Crossreferences: PML:AP000417; PDB:1AB02543.1
 A.Experimental source: cultivar Columbia
 CGenetics:
 A.Map position: 3
 A.Jointons: 3973; 83/2; 111/1; 200/1; 225/2; 252/2; 276/3; 312/2

Query Match 53.0%; Score 939; DB 2; Length 340;
 Best Local Similarity 54.1%; Pred. No. 7.2e-73;
 Matches 167; Conservative 54; Mismatches 91; Indels 6; Gaps 2;

QY 1 MTGSNFKLPPGDISGVFSEFSSQFLVSNSTVSQVLYVPANSLKRYQTGAVL 60
 DD 4 VTPAGKRLNPNFSGDLSLAFPS-NSNDLSSWSKRVLLDVSTSLAGEFLGCAVL 62
 QY 61 DCAVQVPHMSGGLQKMLNDLQENLVGTHGAFRCVCEVSNVWVTSMDQTV 120
 DD 63 DCDHDPGSGVADIKVRYVFKGKEDLGHQKAVCEVSNVWVTSMDQTV 122
 QY 121 KWDPR-----TCNAGTSQPKYTVLSVSGDALVAGRWLVDRVMYQQRRE 175
 DD 123 KWDPRGASQPTQVTLQPRVSVLSVGLHVLVATGRRVMTDLRMVSQQRRE 182
 QY 176 SLVAYCTCTAFKNGKVLSEIGRVAVELDSPFQVKYAFKRLKENNIEQLY 235
 DD 183 SLVAYCTCTCFNGTALSVEGRVAFQFSEVSNVWVTSMDQTV 242
 QY 236 PVNLSFINHTFATGSDGFDNPNKRLQCHRYPTFLASLFSNQTGLVAT 295
 DD 243 PVNLSAFHPTGTFATGCGCFWNMDNNKTKYQVSKYPTFLSALFSNQTGLVAT 302
 QY 296 SVYVMDHTPEHGIFQVTDATKPK 324
 DD 303 STYDEGSKQEPALFVSVEIEVKPK 331

RESULT 2

T72185
 Mitotic checkpoint protein Y49A.6 - Caenorhabditis elegans
 CSpecies: Caenorhabditis elegans

Db 264 MGSLVYVPSIAFYVGVTFVAGDGTFRMDKQNRHL---KQYTLQASIVYCSN 320
 QY 286 NGDTTLALASIVYV---EMDTEIDPDGTFIROVTDATSKP 324
 Db 321 RNSVFAVALSVDMQHGMRNPDYNYIRHATYDEVEK 362
 RESULT 5
 T81301
 Probable mitotic checkpoint WD repeat protein - fission yeast (Schizosaccharomyces pombe)
 CDate: 03-Dec-1999 #sequence_revision 10-Oct-1995 #text_change 26-May-2000
 R:Seaton, J., Churcher, C.W.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.; Hood, V.
 A:Reference number: 221785
 A:Accession: T81301
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Gene: rael
 A:Map position: 1
 A:Cross-references: EMBL:AL021748; PDB:1CA1243.1; GSPDB:GN00066; SPDB:SPAC23H3.08c
 A:Experimental source: strain 972H-; cosmid G23H3
 A:Residues: 1-320 <8KE>
 A:Map position: 1
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 Query Match 29.31; Score 518.5; DB 2; Length 320;
 Best Local Similarity 36.11; Pred. No. 64; 21; Indels 21; Gaps 4;
 Matches 119; Conservative 57; Mismatches 111; Indels 21; Gaps 4;
 QY 14 EDGISSEKSPNTSOFILVSSMDYVRYDVPANRML-NQRTGAVLDCQVDFTHMS 72
 Db 12 RGDSLSVITFSSVNNELASQDGSILRFQISENPELAKVLSPEYDTEKTLV 71
 QY 73 GDHLQKHLMDLNDENLQVHDPKICVYCYCEVMMVVGSMQTVKMDPRTPCKAG 132
 Db 72 QDLGDTVTTLQDNLNHFQVQNGEKVSGCISKLEKCTISBNSKSFVNVGVN---- 126
 QY 133 TFSQ-----EKVYTLVSSGDLVGTAGRVVMDLRNMDVQVQVRRSSLKQTRCI 185
 Db 127 ---KQVPEQGTGKTPASSNSDNLVGCSEKENVYDNLKLPQRRFSKFTVIR 184
 QY 186 RAFPNGQVLLSISGRVAVETLDPSPVQKQYAEKCEKLENNTEIQIVVPAISHMI 245
 Db 195 CONQNFATVSSSILRSTVITINRQASKNFTFALQKQVQVIVVMDLKHNI 244
 QY 246 NHTPATGSDQFNVMPFNKGLQCPHRYTSLASLAFSGDTTLATASYVMEMDTE 305
 Db 245 KHTLALAGQGVNMDIQVRLVAFENKISLSFNVDSKVNALQCAQ-----E 298
 QY 306 HPEDGIFIRVQVTDATSKP 324
 Db 299 EAGQNIYVHALESNPAKP 317
 RESULT 6
 A56119
 RNA export protein rael - fission yeast (Schizosaccharomyces pombe)
 N:Alternative names: poly(a)-rna export protein
 C:Species: Schizosaccharomyces pombe
 C:Dates: 18-Oct-1999 #sequence_revision 19-Oct-1995 #text_change 26-May-2000
 C:Accession: A56119; T39547
 S:Brown, J.A.; Bhattacharya, A.; Ghosh, A.; Whalen, W.; Fitzgerald, E.; Dhar, R.
 A:Ref. Biol. Chem. 270, 7411-7420, 1995
 A:Reference number: A56119; NUID:95221400; PMID:7706287
 A:Accession: A56119
 A:Status: preliminary
 A:Gene: rael
 A:Residues: 1-352 <8RO>
 A:Cross-references: GB:U14951; NID:9625093; PDB:AA8631.1; PDB:9625094

R:Rejndream, M.A.; Barrall, B.G.; Beck, A.; Reinhardt, R.
 A:Reference number: 221862
 A:Accession: T39547
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Gene: rael
 A:Map position: 2
 A:Cross-references: EMBL:AL021748; PDB:1CA16856.1; GSPDB:GN00067; SPDB:SPBCE16A3.05c
 A:Experimental source: strain 972H-; cosmid c16A3
 A:Residues: 1-352 <8RO>
 A:Map position: 2
 A:Introns: 24/3; 50/2
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 Query Match 28.91; Score 511.5; DB 2; Length 352;
 Best Local Similarity 36.11; Pred. No. 70; 21; Indels 15; Gaps 7;
 Matches 119; Conservative 58; Mismatches 142; Indels 15; Gaps 7;
 QY 6 EFKLNQPDGIESVSPNTSOFILVSSMDYVRYDVPANRML-NRKYQVHGVAVLC 62
 Db 22 DYEVAQPDGIESVSPNTSOFILVSSMDYVRYDVPANRML-NRKYQVHGVAVLC 60
 QY 63 AF-YDPTIAMSGLGDLQHLQDLNLTQDNLQVHDPKICVYCYCEV---VMVTSQD 118
 Db 81 NMDSDTQVSSGVSINKAKVFDIQTCQVQVAVDAVACVTRVEAGTSPFKLANSP 140
 QY 119 TVKMDPRTPCKAGTFIRVQVTDATSKP 324
 Db 141 TLKLVQSGTPIATVSLPRTVYVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 200
 QY 179 KYOTCRIRAFNPKQGVSSISGRVAVETLDPSPVQKQYAEKCEKLENNTEIQIV 238
 Db 201 KFTNSLACFINGSTVAGVGRGALQNTDEKNSQ--NESPCHBNQAGNSADYSVM 258
 QY 239 AISHNHNHTPATGSDQFNVMPFNKGLQCPHRYTSLASLAFSGDTTLATASY-- 297
 Db 259 SINKHQVGTGVSQVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 318
 QY 298 ---NEMDQDTEHPEDGIFIRVQVTDATSKP 324
 Db 319 WKQYTPNNAQLP-AKTMHPVPPQDEIHP 347
 RESULT 7
 T39773
 Hypothetical protein P108.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Dates: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000
 S:Baaham, V.
 Submitted to the EMBL Data Library, September 1996
 A:Reference number: 219315
 A:Accession: T39773
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 373 <8RO>
 A:Cross-references: EMBL:280316; PDB:1CA02280.1; GSPDB:GN00019; CESP:P108.3
 A:Gene: CESP:P108.3
 A:Experimental source: clone F108
 A:Introns: 36/3; 67/3; 345/3
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 Query Match 32.61; Score 418.5; DB 2; Length 373;
 Best Local Similarity 31.51; Pred. No. 3; 8; 28;
 Matches 109; Conservative 63; Mismatches 145; Indels 29; Gaps 11;
 QY 2 TGSNHLKQK 56
 Db 30 TQNDLVDVQAPEDTQVTFESFTPDQKPLACGSDNQTITWFMFNDATFEGQAQNI 89

R; Kaneo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. 2001. Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena cylindrica*. *Genome Research* 11:2595-2598; PMID:11759840
 A:Accession: AF12155
 A:Status: preliminary
 A:Title: preliminary
 A:Residues: 11258 <KUR>
 A:Cross-references: GB:BA000019; PID:GB77459.1; PID:g17131893; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 A:Gene: alr-2800

Query Match: 11.8%; Score 209; DB 2; Length 1258;
 Best Local Similarity 24.5%; Pred. No. 1.9e-09;
 Matches 48; Conservative 60; Mismatches 139; Indels 42; Gaps 11;
 Qy 15 DGISSVSPNTSOFVLVSSVTSVRLVYDV-PANSMRLKVTQTFGLVLDCAF-YDTHANS 72
 Db 765 DVNVCVAFSP-GNTLSSANHTLMDVSGQCKLTLSTHVSFVSADQTLAS 827
 Qy 73 GLDHLQHLMDLNTDQ-ENLVGTADHAPICVEYCEVNVYTSQDTVKLMDRTP-C 129
 Db 828 GSDNLTGTLTQDQ-TLVNVTGTLTQDQ-TLVNVTGTLTQDQ-TLVNVTGTLTQDQ 886
 Qy 130 NAGTFEQEKVNTVTS-GORLVGTAGRVVLNDELN- 171
 Db 897 IYTLKHNHNVCFVSFVSGDTGLACVSLDQVSLVLMCTGCLKAWYQNTMALVAFSP 946
 Qy 172 QRESELKVTCTIRAFKPGKGVLSIEGRVAVEV-LDSEFVQKKVAF- 221
 Db 947 DROLSSNKNYKVLMDQVTKGYSLEIGHDFITGLASFSQDTLASESTSSVLWN 1006
 Qy 222 ----KCHRLKENNIQVFNALSFNTHTFATGSGSGGVNTWDPENKELQPHRYPT 277
 Db 1007 ISTQCFQILLHEDMWT----AVVHPQKQIATGSACTVKLMIISTQGLATLSHSD 1063
 Qy 278 STASLAFNSGDTLALAS 296
 Db 1064 KILGWASFPQGLLASASA 1082

RESULT 14
 AE1810
 A:Accession: AF12155
 A:Status: preliminary
 A:Title: Nostoc protein (imported) - Nostoc sp. (strain PCC 7120)
 A:Cross-references: GB:BA000019; PID:BA077555.1; PID:g171315007; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 A:Gene: alr-2800
 Query Match: 11.4%; Score 206; DB 2; Length 1227;
 Best Local Similarity 24.8%; Pred. No. 3.3e-09;
 Matches 84; Conservative 61; Mismatches 132; Indels 62; Gaps 15;
 Qy 17 ISSVSPNTSOFVLVSSVTSVRLVYDV-PANSMRLKVTQTFGLVLDCAF-YDTHANS 75
 Db 903 VYSVAFSP-SQTLASORDQYDTIGMLTKTGTCPIGLGQGRIRAFVAFPOKTLASSA 961
 Qy 76 DROLKHEL-NTDQENLVGT----HDAPICVEYCEVNVYTSQDTVKLMDRTP-PCH 130

Db 961 DNTGLKMDISDTHSKVIGLITVCTHTNVVTVVWVSPDQTLASSKNTIRLMDKQDGL 1021
 Qy 1222 AKTQSGPEKVTYKSSQKTLV-OTAGRRVWDLWRHMGVVOORRESLKYVTCIRAF 188
 Db 1033 QGLGHSHHWTVVAFSPGRTLASGASKEIXINDV- 1066
 Qy 189 PKHQGVLS----SIGER-VAVEYDLSPEVQKKVAFKCHLKENNTEQIYFNALISFN 244
 Db 1067 TDPQGVNSVAFSLQTLTASQDQVTLKWLATGECVTLKQHE-KNY- -SVAFSP 1122
 Qy 245 INTFTATGSGGVNTWDPENKELQCF-HRYFTSTIASLAFSNDGTTLA- 292
 Db 1123 NQGLASSEDVTVLMDISTGSCVOTLASHHTAFVAFSPDRLSSDSEKIQUM 1182
 Qy 293 -----TASSWTE-MDUTEHPDGIFIVQVDAE 320
 Db 1183 DMQKSRKLTGLFALFENMDYID-----ITQIDNE 1214
 RESULT 15
 F782.18 protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #test_change 02-Nov-2001
 A:Accession: H62392
 A:Status: preliminary
 A:Title: F782.18 protein
 A:Cross-references: GB:BA000019; PID:BA077555.1; PID:g171315007; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 A:Gene: alr-2800
 Query Match: 11.2%; Score 198.5; DB 2; Length 140;
 Best Local Similarity 35.7%; Pred. No. 7.9e-10;
 Matches 45; Conservative 26; Mismatches 46; Indels 9; Gaps 4;
 Qy 4 SNFKLNPQEDISSVFNQNTSVLSHSDVRLVYDV- 57
 Db 16 INNSTEITPATDTSISLSISFSFK-ADVLVYSDVQVCHETSTDSISASPKSVNSHSD 74
 Qy 58 AVLDCAFYDP-THAMSGDLQHLMD-LNTDQENLVGTADHAPICVEYCEVNVYTSQ 115
 Db 75 PVLCSAMDDQDITVFGQDKQKQWMLPSLQSGRTSVNDFAFNQIATIRGQNLVATGS 134
 Qy 116 MDQVTK 121
 Db 135 MDKTLK 140

Search completed: November 12, 2003, 20:10:46
 Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein protein search, using sw model

Run on: November 12, 2003, 20:05:12 / Search time 43 Seconds

(without alignments)
1210.752 Million cells updates/sec

Title: US-10-084-700-2

Perfect score: 1772

Sequence: 1 MTSSEFLNKPDSIGSV.....DGIFRQYDTRHPSCT 328

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 3. Genesep 13Jun03.*
1: /SIDSL/gcgdata/genesep/genesep-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/genesep/genesep-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/genesep/genesep-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/genesep/genesep-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/genesep/genesep-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/genesep/genesep-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/genesep/genesep-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/genesep/genesep-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/genesep/genesep-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/genesep/genesep-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/genesep/genesep-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/genesep/genesep-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/genesep/genesep-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/genesep/genesep-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/genesep/genesep-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/genesep/genesep-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/genesep/genesep-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/genesep/genesep-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/genesep/genesep-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/genesep/genesep-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/genesep/genesep-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/genesep/genesep-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/genesep/genesep-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/genesep/genesep-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1772	100.0	328	20	AA196288
2	1772	100.0	328	21	AA196289
3	1772	100.0	328	22	AA196290
4	1772	100.0	328	23	AA196291
5	1763	99.5	328	24	AA196292
6	10539	53.0	340	21	AA196293
7	10539	53.0	340	22	AA196294
8	705.5	39.8	166	23	AA196295
9	591.5	32.8	343	21	AA196296

10	579	32.7	191	21	AA196095
11	578.5	32.6	349	21	AA196096
12	578.5	32.6	349	21	AA196097
13	571	32.2	107	20	AA196133
14	571	32.2	107	20	AA196134
15	520.5	28.7	368	20	AA196147
16	509	28.7	413	21	AA196174
17	481	27.1	231	21	AA196216
18	481	27.1	231	21	AA196217
19	480	27.1	232	21	AA196218
20	475	26.8	238	21	AA196219
21	475	26.8	238	21	AA196220
22	475	26.8	238	21	AA196221
23	475	26.8	238	21	AA196222
24	381	21.5	535	22	AA196223
25	356	20.1	341	24	AA196224
26	356	20.1	341	24	AA196225
27	356	20.1	341	24	AA196226
28	258	14.6	59	20	AA196227
29	255	14.4	158	21	AA196228
30	202.5	12.6	158	22	AA196229
31	202.5	12.6	158	22	AA196230
32	197.5	11.1	100	21	AA196231
33	196	11.1	388	22	AA196232
34	185	10.5	497	22	AA196233
35	185	10.5	497	22	AA196234
36	182.5	10.3	357	22	AA196235
37	177	10.0	310	21	AA196236
38	177	10.0	310	21	AA196237
39	177	10.0	310	21	AA196238
40	176.5	10.0	375	23	AA196239
41	176.5	10.0	375	23	AA196240
42	176.5	10.0	375	24	AA196241
43	176.5	10.0	375	24	AA196242
44	173.5	9.9	433	21	AA196243
45	173.5	9.9	486	21	AA196244

ALIGNMENTS

RESULT 1
AA196288
XX AA196288 standard; Protein; 328 AA.

XX AA196288;

DT DT 23-ANG-1999 (first entry)

XX XX Human BUB3, a substrate for BUB1A and BUB1B.

XX BUB3; human; protein kinase; checkpoint control; mitosis;

KW kinase; cell proliferation; cancer; tumour; therapy.

XX Homo sapiens.

XX W0928334-A1.

PD 10-JUN-1999.

XX 01-DEC-1998; 98MO-025415.

XX 01-DEC-1997; 97US-0067033.

XX (FOXC-) FOX CHARGE CANCER CENT.

XX Chan G, Jablonski S, Yen T;

XX WPI; 1999-371094/31.

XX R-PS08; AA196292.

XX Human ovarian anti

XX Human BUB3 and proteins involved in mitotic checkpoint control

PR	21-JUL-1999	99US-01450088
PR	21-JUL-1999	99US-01450088
PR	22-JUL-1999	99US-01450085
PR	22-JUL-1999	99US-01450085
PR	22-JUL-1999	99US-01451532
PR	22-JUL-1999	99US-01451532
PR	23-JUL-1999	99US-01451532
PR	23-JUL-1999	99US-01451532
PR	23-JUL-1999	99US-01452218
PR	23-JUL-1999	99US-01452218
PR	23-JUL-1999	99US-01452766
PR	23-JUL-1999	99US-01452766
PR	27-JUL-1999	99US-01455918
PR	27-JUL-1999	99US-01455918
PR	27-JUL-1999	99US-01460336
PR	27-JUL-1999	99US-01460336
PR	02-AUG-1999	99US-01463869
PR	02-AUG-1999	99US-01463869
PR	03-AUG-1999	99US-01470308
PR	03-AUG-1999	99US-01470308
PR	04-AUG-1999	99US-01474031
PR	04-AUG-1999	99US-01474031
PR	05-AUG-1999	99US-01474933
PR	05-AUG-1999	99US-01474933
PR	06-AUG-1999	99US-01479375
PR	06-AUG-1999	99US-01479375
PR	09-AUG-1999	99US-01481731
PR	09-AUG-1999	99US-01481731
PR	12-AUG-1999	99US-01483431
PR	12-AUG-1999	99US-01483431
PR	13-AUG-1999	99US-01485465
PR	13-AUG-1999	99US-01485465
PR	13-AUG-1999	99US-01486864
PR	13-AUG-1999	99US-01486864
PR	16-AUG-1999	99US-01494522
PR	16-AUG-1999	99US-01494522
PR	18-AUG-1999	99US-01494522
PR	20-AUG-1999	99US-01497223
PR	20-AUG-1999	99US-01497223
PR	20-AUG-1999	99US-01499302
PR	20-AUG-1999	99US-01499302
PR	23-AUG-1999	99US-01505636
PR	23-AUG-1999	99US-01505636
PR	25-AUG-1999	99US-01508844
PR	26-AUG-1999	99US-01510664
PR	27-AUG-1999	99US-01510664
PR	27-AUG-1999	99US-01510880
PR	30-AUG-1999	99US-01531303
PR	30-AUG-1999	99US-01531303
PR	01-SEP-1999	99US-01534338
PR	01-SEP-1999	99US-01534338
PR	07-SEP-1999	99US-01532363
PR	07-SEP-1999	99US-01532363
PR	10-SEP-1999	99US-01530701
PR	10-SEP-1999	99US-01530701
PR	13-SEP-1999	99US-01537578
PR	13-SEP-1999	99US-01537578
PR	16-SEP-1999	99US-01540439
PR	16-SEP-1999	99US-01540439
PR	20-SEP-1999	99US-01544779
PR	20-SEP-1999	99US-01544779
PR	22-SEP-1999	99US-01551339
PR	22-SEP-1999	99US-01551339
PR	23-SEP-1999	99US-01554586
PR	23-SEP-1999	99US-01554586
PR	28-SEP-1999	99US-01556596
PR	28-SEP-1999	99US-01556596
PR	04-OCT-1999	99US-01557117
PR	04-OCT-1999	99US-01557117
PR	06-OCT-1999	99US-01560029
PR	06-OCT-1999	99US-01560029
PR	07-OCT-1999	99US-01580293
PR	07-OCT-1999	99US-01580293
PR	08-OCT-1999	99US-01582832
PR	08-OCT-1999	99US-01582832
PR	12-OCT-1999	99US-01585669
PR	12-OCT-1999	99US-01585669
PR	13-OCT-1999	99US-01587284
PR	13-OCT-1999	99US-01587284
PR	13-OCT-1999	99US-01592525
PR	13-OCT-1999	99US-01592525
PR	14-OCT-1999	99US-01593330
PR	14-OCT-1999	99US-01593330
PR	14-OCT-1999	99US-01596637
PR	14-OCT-1999	99US-01596637
PR	18-OCT-1999	99US-01595864
PR	18-OCT-1999	99US-01595864

[illegible]

PR 01-JUN-1999; 99US-011728.
 PR 02-JUN-1999; 99US-011729.
 PR 07-JUN-1999; 99US-011724.
 PR 08-JUN-1999; 99US-0118094.
 PR 09-JUN-1999; 99US-0118540.
 PR 10-JUN-1999; 99US-0118540.
 PR 14-JUN-1999; 99US-0119119.
 PR 16-JUN-1999; 99US-0119452.
 PR 16-JUN-1999; 99US-0119453.
 PR 16-JUN-1999; 99US-0119453.
 PR 18-JUN-1999; 99US-0119454.
 PR 18-JUN-1999; 99US-0119455.
 PR 18-JUN-1999; 99US-0119456.
 PR 18-JUN-1999; 99US-0119457.
 PR 18-JUN-1999; 99US-0119458.
 PR 18-JUN-1999; 99US-0119459.
 PR 18-JUN-1999; 99US-0119460.
 PR 18-JUN-1999; 99US-0119461.
 PR 18-JUN-1999; 99US-0119462.
 PR 18-JUN-1999; 99US-0119463.
 PR 18-JUN-1999; 99US-0119464.
 PR 21-JUN-1999; 99US-0119817.
 PR 22-JUN-1999; 99US-0119889.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 23-JUN-1999; 99US-0140355.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140921.
 PR 30-JUN-1999; 99US-0141099.
 PR 01-JUL-1999; 99US-0111842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 06-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142820.
 PR 12-JUL-1999; 99US-0142977.
 PR 12-JUL-1999; 99US-0143542.
 PR 13-JUL-1999; 99US-0143542.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145313.
 PR 27-JUL-1999; 99US-0145313.
 PR 27-JUL-1999; 99US-0145319.
 PR 28-JUL-1999; 99US-0145951.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146186.
 PR 02-AUG-1999; 99US-0146186.
 PR 02-AUG-1999; 99US-0146189.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147192.

PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147301.
 PR 06-AUG-1999; 99US-0147301.
 PR 09-AUG-1999; 99US-0147423.
 PR 09-AUG-1999; 99US-0147423.
 PR 11-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148565.
 PR 16-AUG-1999; 99US-0149175.
 PR 16-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149456.
 PR 18-AUG-1999; 99US-0149456.
 PR 20-AUG-1999; 99US-0149733.
 PR 20-AUG-1999; 99US-0149733.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149902.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150684.
 PR 27-AUG-1999; 99US-0151063.
 PR 27-AUG-1999; 99US-0151063.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 31-AUG-1999; 99US-0151438.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155466.
 PR 23-SEP-1999; 99US-0155466.
 PR 26-SEP-1999; 99US-0156158.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 04-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157855.
 PR 07-OCT-1999; 99US-0158039.
 PR 08-OCT-1999; 99US-0158222.
 PR 11-OCT-1999; 99US-0158222.
 PR 13-OCT-1999; 99US-0159291.
 PR 13-OCT-1999; 99US-0159291.
 PR 13-OCT-1999; 99US-0159291.
 PR 14-OCT-1999; 99US-0159120.
 PR 14-OCT-1999; 99US-0159120.
 PR 14-OCT-1999; 99US-0159311.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159637.
 PR 16-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161993.
 PR 28-OCT-1999; 99US-0161993.
 PR 28-OCT-1999; 99US-0162142.
 PR 28-OCT-1999; 99US-0162142.

Query Match

32.6%; Score 570.5; DB 21; Length 349;

Best Local Similarity 38.7%; Pred. No. 5.8e-52;
Matches 129; Conservative 60; Mismatches 125; Indels 19; Gaps 8;
Qy 4 SHEKMGKDEGDIKVESNAGSOGRLVSWFSYRLIDVDP-----ASNMRLKVGHTS 73
15 NKSVEYTPSPALSSISLSFR-ADLVATSDMNOVCHLSRGASLASAPKASISHDQ 57
Qy 56 AVLCQKVP-THANGSLGHLQKMD-LATQDNLVGHDAIPICVEYCPISVNNWVTS 115
74 PVLCAKRDGDTGVSSGGCDKAKWLLSGQGVTVANHGSLIAWVPGKLLATIS 133
Qy 116 WNTVYLDNPTFCNMGTSFGSPKPVYLVSSGDLGLVGTAGKRVLWDLRMGVQORE 175
134 NQTLKAVDQKQVHGTQQLPDKCTLLSVSLVPLVADENLVNGLVSNPOTFEKID 193
Qy 176 SLKQITICIFAPFMKQGVLSSEYQVAVELDPSPQKGLGCFHARTFSLASFGNQTLLAS 235
236 PVNLSFHNIHNTATGSDQVNTWPPKMKGLGCFHARTFSLASFGNQTLLAS 295
248 SVNLSFHPVGTATGSDQVNTWPPKMKGLGCFHARTFSLASFGNQTLLAS 307
Qy 286 SYWKEGDTHT-----PDGIGITROVDAETKPK 324
308 CYDNGKGAENHPATKSSIFILNLPQESVAK 340

RESULT 12
ID AAG45336 standard; Protein; 349 AA.
XX AAG45336;
XX AAG45336;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 56906.
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999;
XX 05-MAR-1999;
XX 09-MAR-1999;
XX 23-MAR-1999;
XX 23-MAR-1999;
XX 23-MAR-1999;
XX 01-APR-1999;
XX 06-APR-1999;
XX 08-APR-1999;
XX 19-APR-1999;
XX 21-APR-1999;
XX 21-APR-1999;
XX 28-APR-1999;
XX 30-APR-1999;
XX 04-MAY-1999;
XX 06-MAY-1999;
XX 06-MAY-1999;
XX 07-MAY-1999;
XX 11-MAY-1999;

14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
19-MAY-1999;
20-MAY-1999;
21-MAY-1999;
21-MAY-1999;
25-MAY-1999;
27-MAY-1999;
28-MAY-1999;
01-JUN-1999;
01-JUN-1999;
04-JUN-1999;
07-JUN-1999;
08-JUN-1999;
08-JUN-1999;
10-JUN-1999;
14-JUN-1999;
16-JUN-1999;
17-JUN-1999;
17-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
23-JUN-1999;
23-JUN-1999;
23-JUN-1999;
28-JUN-1999;
29-JUN-1999;
30-JUN-1999;
01-JUL-1999;
02-JUL-1999;
06-JUL-1999;
08-JUL-1999;
08-JUL-1999;
12-JUL-1999;
13-JUL-1999;
14-JUL-1999;
16-JUL-1999;
16-JUL-1999;
19-JUL-1999;
19-JUL-1999;
19-JUL-1999;
19-JUL-1999;
19-JUL-1999;
20-JUL-1999;
20-JUL-1999;
20-JUL-1999;
20-JUL-1999;
21-JUL-1999;
21-JUL-1999;
21-JUL-1999;
22-JUL-1999;
22-JUL-1999;
23-JUL-1999;
23-JUL-1999;
23-JUL-1999;

99US-0134218;
99US-0134219;
99US-0134221;
99US-0134370;
99US-0134948;
99US-0134948;
99US-0135124;
99US-0135353;
99US-0135359;
99US-0136299;
99US-0136392;
99US-0136782;
99US-0137222;
99US-0137502;
99US-0137724;
99US-0138094;
99US-0138094;
99US-0138494;
99US-0139119;
99US-0139452;
99US-0139452;
99US-0139453;
99US-0139454;
99US-0139454;
99US-0139455;
99US-0139456;
99US-0139456;
99US-0139458;
99US-0139458;
99US-0139460;
99US-0139461;
99US-0139462;
99US-0139463;
99US-0139750;
99US-0139751;
99US-0139763;
99US-0139899;
99US-0139899;
99US-0140353;
99US-0140354;
99US-0140823;
99US-0140823;
99US-0140891;
99US-0141287;
99US-0141287;
99US-0142554;
99US-0142555;
99US-0142590;
99US-0142590;
99US-0142800;
99US-0142877;
99US-0142877;
99US-0143542;
99US-0143624;
99US-0143624;
99US-0144086;
99US-0144086;
99US-0144225;
99US-0144225;
99US-0144331;
99US-0144332;
99US-0144334;
99US-0144334;
99US-0144335;
99US-0144632;
99US-0144632;
99US-0144884;
99US-0144884;
99US-0145088;
99US-0145088;
99US-0145088;
99US-0145087;
99US-0145089;
99US-0145145;
99US-0145218;

DR N-PSDB; AAX41166.
 XX New nucleic acid encoding human secreted proteins - obtained from
 XX cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
 XX kidney, lung, umbilical cord, placenta and colon tissue
 XX Claim 27, Page 697, 824pp; English.
 XX
 XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
 XX human secreted proteins, and encode the proteins given in AAX41262 to
 XX AAX41347. The proteins are secreted from the cells of the placenta
 XX and an N-terminal fragment of a secreted protein. The nucleic acid
 XX sequences can be used for producing secreted human gene products. They
 XX can also be used to develop products for diagnosis and therapy. The
 XX sequences can also be used for obtaining corresponding promoters for
 XX proliferation/differentiation activity, haematopoiesis regulating
 XX activity, tissue growth regulating activity, reproductive hormone
 XX regulating activity, chemotactic activity, haemostatic and
 XX activity, tumour inhibition activity or other activities. The products
 XX can be used in forensic, gene therapy and chromosome mapping procedures.
 XX The sequences can also be used for obtaining corresponding promoters for
 XX directing extracellular secretion of a polypeptide or the insertion of a
 XX polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 XX Sequence 107 AA;
 XX
 XX Query Match 32-28; Score 571; DB 20; Length 107;
 XX Best Local Similarity 98.11; Pred. No. 68-52; Indels 0; Gaps 0;
 XX Matches 104; Conservative 0; Mismatches 2;
 XX
 XX Qy 1 MTGSENFPLKPPEDGIGSSVFESVNTSQTLLVSSMTSVRLVDPANSMRLKYQTHAVL 60
 XX Db 1 MTGSENFPLKPPEDGIGSSVFESVNTSQTLLVSSMTSVRLVDPANSMRLKYQTHAVL 60
 XX
 XX Qy 61 DCAFYPTTHANGSGGLDQLKHLNDQENLVGTHDAPICVCEVC 106
 XX Db 61 DCAFYPTTHANGSGGLDQLKHLNDQENLVGTHDAPICVCEVC 106
 XX
 XX RESULT 14
 XX ABB64775 32-28; Score 571; DB 20; Length 107;
 XX ABB64713 standard; Protein: 346 AA.
 XX AC ABB64713;
 XX NT 26-MAR-2002 (first entry)
 XX
 XX XX Drosophila melanogaster polypeptide SEQ ID NO 20931.
 XX XX Drosophila; developmental biology; cell signalling; insecticide;
 XX XX pharmaceutical.
 XX XX
 XX XX Drosophila melanogaster.
 XX XX W0200171042-42.
 XX XX
 XX XX 27-SEP-2001.
 XX XX
 XX XX 23-MAR-2001; 2001MO-US093231.
 XX XX
 XX XX 23-MAR-2000; 2000US-1914379.
 XX XX 11-JULY-2000; 2000US-0611150.
 XX XX (PEPE) PE CORP NY.
 XX XX
 XX XX Venter JC, Adams M, Li PHD, Myers BW,
 XX XX WPT; 2001-656860/75.
 XX XX N-PSDB; ABL08816.
 XX XX
 XX XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 XX
 XX Disclosure: SEQ ID NO 20931; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX cell-cell interactions in higher eukaryotes for the development of
 XX diseases genomic DNA sequences (AB161616-AB161617), expressed DNA
 XX sequences (AB01840-AB16175) and the encoded proteins
 XX (AB057737-AB07272).
 XX For this patent did not form part of the printed
 XX specification, and was obtained in electronic format directly from WIPO
 XX at http://wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 346 AA;
 XX
 XX Query Match 29-44; Score 521.5; DB 22; Length 346;
 XX Best Local Similarity 36.44; Pred. No. 5.86-46;
 XX Matches 123; Conservative 59; Mismatches 133; Indels 23; Gaps 11;
 XX
 XX Qy 5 NBYKLVNPPEDGIGSSVFESVNTSQTLLVSSMTSVRLVDPANSMRLKYQTHAVL 59
 XX Db 12 NBYKLVNPPEDGIGSSVFESVNTSQTLLVSSMTSVRLVDPANSMRLKYQTHAVL 59
 XX
 XX Qy 60 LD-CAPFYPTTHANGSGGLDQLKHLNDQENLVGTHDAPICVCEVC - PEVNVVTGSM 116
 XX Db 72 LDVMSDGSQVASCQKYLKGLASQVQVAAHGVPTCHWGWGPTTCMLTGM 131
 XX
 XX Qy 117 DQYVNDPRTFNAQTSQPEKVTYLSQSDRLVGTAGRWLWDLNMGVQQRSS 176
 XX Db 132 DXTLQWDSFSPMPTNIPERCYCADVEPMVAVTARGLLIYLSQSPETKQES 191
 XX
 XX Qy 177 SLKVTGTCIAFAFNKO---GVLSSTEGRVAVELDPSVEQVKYAFKRLK-ENNI 231
 XX Db 192 PLKYQHAISIFRDKKEPTGALSGTEGVIAQVNFGR--PKDMFTFCHRTTGTST 249
 XX
 XX Qy 232 SQLYPVAHLSFHHHTMTGSDGFVNDPNKELCOFHRYPTSIALSFGNDGTTL 291
 XX Db 250 ODIVAVNDIAFPHVHTLVTVGSDGDTSPMDKARTKLKSSSTWDQSTKCGFNANQIF 309
 XX
 XX Qy 292 ALASAEVYE---NDDTEPDQTFIRQVTDATETPK 324
 XX Db 310 NAYGVDSNGKHIEFTPAKFKQ--IFLRSCYD-ELAKPR 344
 XX
 XX RESULT 15
 XX AAY31647
 XX ID AAY31647 standard; Protein: 368 AA.
 XX AC AAY31647;
 XX NT 02-NOV-1999 (first entry)
 XX XX
 XX XX Human transport-associated protein-9 (TRANP-9).
 XX XX Transport-associated protein; TRANP; nuclear pore; nuclear transport;
 XX XX hypercholesterolaemia; diagnosis; treatment.
 XX XX Homo sapiens.
 XX XX
 XX XX Location/Qualifiers
 XX XX 101..115
 XX XX Region
 XX XX FT Region
 XX XX Note="Beta transducin family Ttp-Aap repeat"
 XX XX /note="Beta transducin family Ttp-Aap repeat"
 XX XX MO941373-A2.
 XX XX 19-NOV-1999.

Result No.	Score	Query Match	Length	ID	Description
1	100.0	59	4	US-09-025-6811	Sequence 17, Appl
2	226	59	4	US-09-025-1155	Sequence 17, Appl
3	175.5	11	251	US-09-251-100A-13	Sequence 13, Appl
4	175.5	10.0	275	US-09-063-713-13	Sequence 13, Appl
5	175.5	10.0	375	US-09-580-500-12	Sequence 12, Appl
6	176.5	10.0	375	US-09-580-500-12	Sequence 12, Appl
7	190.5	9.6	350	US-09-231-147A-12	Sequence 12, Appl
8	190.5	9.6	350	US-09-231-147A-12	Sequence 12, Appl
9	191.5	9.6	340	US-08-150-8021-38	Sequence 38, Appl
10	191.5	9.6	340	US-08-150-8021-38	Sequence 38, Appl
11	192.5	9.6	340	US-08-477-346-33	Sequence 33, Appl
12	192.5	9.6	340	US-08-477-346-33	Sequence 33, Appl
13	193.5	9.6	340	US-09-245-019-18	Sequence 18, Appl
14	193.5	9.6	340	US-09-245-019-18	Sequence 18, Appl
15	194.5	9.6	704	US-08-168-582-18	Sequence 18, Appl
16	194.5	9.6	704	US-08-168-582-18	Sequence 18, Appl
17	196	9.4	318	US-08-190-8021-33	Sequence 33, Appl
18	196	9.4	318	US-08-477-346-33	Sequence 33, Appl
19	196	9.4	318	US-08-477-346-33	Sequence 33, Appl
20	197	9.2	715	US-08-150-8021-63	Sequence 63, Appl
21	197	9.2	715	US-08-150-8021-63	Sequence 63, Appl
22	198	9.2	713	US-08-477-346-33	Sequence 33, Appl
23	198	9.2	713	US-08-477-346-33	Sequence 33, Appl
24	198	9.2	713	US-09-166-887-33	Sequence 33, Appl
25	198	9.2	713	US-09-166-887-33	Sequence 33, Appl
26	199.5	9.1	340	US-08-190-8021-42	Sequence 42, Appl
27	199.5	9.1	340	US-08-190-8021-42	Sequence 42, Appl
28	199.5	9.1	340	US-08-477-346-33	Sequence 33, Appl

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```

APPLICANT: Lal, Preeti
APPLICANT: Guegler, Neil J.
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 6
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: 08-Jun-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Carrone, Michael C
REGISTRATION NUMBER: 39,132
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
LENGTH: 375 amino acids
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LINGNOT01
LIBRARY: 059553
US-09-063-743-1

```

```

Query Match 10.0%; Score 176.5; DB 3; Length 375;
Best Local Similarity 25.8%; Pred. No. 8.2e-11;
Matches 80; Conservative 51; Mismatches 142; Indels 37; Gaps 13;

QY 14 EDGSSVKESFNSTQFSLVSWDSVRLVDYPA---NSMRKLYQITGAVLDCAE---YDTH 69
DB 66 EGVYCKEHNHST-LASAGFDRLILLVYGGCDNATLK-GHSGAWELWITGSM 123
QY 70 AMSGGLDHLQDHLDTQQ--ENLVGTHDAPFRCVEYCEVNVVYTGSHDQTVKLMDET 127
DB 124 LFSASTUKTVAWDSGEVRLKAGHSTFVNSCTPARGPQVCTSDGTGKLMDFK 183
QY 128 PNAGTFTSPQPKY--YTLVSQDRLVGTAGRLVLDLRNMGTVQGRSSSLAYQTR-- 183
DB 184 KAAIQTONTQYLVNTSDQIIISGINDIKWDL-----RKNLLYTMGH 234
QY 184 ---CTAEPKQGVLSIEGRVAVELDPSPVQKKYAFKRLKENNIEQIYVPA 239
DB 235 ADVSTGLSSGGSTLLSAMON--TVVWDVFPAPKE---RCVKTFQGNVH--FENL 288
QY 240 I--SPHNHNTATGSDQFVWIDPKNKLCQPHRYPTASLAFNSDGTLLAASS- 296
DB 289 LACSNSPDGSKIAAGSDRFFVYMDTSSRLYKLPGLHGSINVFAPHDPTIIASSD 348
QY 297 ---YKWDG 303
DB 349 XRLVMEFSE 358

```

```

RESULT 6
US-09-590-540-1
Sequence 1, Application US/09590540
FASTA FILE: 09590540.faa
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Lal, Preeti
Guegler, Neil J.
Patterson, Chandra
Corley, Neil C.
ADDRESS: INCYTE GENOMICS, INC.
STREET: 3160 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 3.0
CURRENT APPLICATION DATA: US/09/590,540
FILING DATE: 08-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA: 09/063,743
REGISTRATION NUMBER: 39,132
TELECOMMUNICATION INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: David C. Streeter
REGISTRATION NUMBER: 468
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LINGNOT01
LIBRARY: 059553
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
Query Match 10.0%; Score 176.5; DB 4; Length 375;
Best Local Similarity 25.8%; Pred. No. 8.2e-11;
Matches 80; Conservative 51; Mismatches 142; Indels 37; Gaps 13;

QY 14 EDGSSVKESFNSTQFSLVSWDSVRLVDYPA---NSMRKLYQITGAVLDCAE---YDTH 69
DB 66 EGVYCKEHNHST-LASAGFDRLILLVYGGCDNATLK-GHSGAWELWITGSM 123
QY 70 AMSGGLDHLQDHLDTQQ--ENLVGTHDAPFRCVEYCEVNVVYTGSHDQTVKLMDET 127
DB 124 LFSASTUKTVAWDSGEVRLKAGHSTFVNSCTPARGPQVCTSDGTGKLMDFK 183
QY 128 PNAGTFTSPQPKY--YTLVSQDRLVGTAGRLVLDLRNMGTVQGRSSSLAYQTR-- 183
DB 184 KAAIQTONTQYLVNTSDQIIISGINDIKWDL-----RKNLLYTMGH 234
QY 184 ---CTAEPKQGVLSIEGRVAVELDPSPVQKKYAFKRLKENNIEQIYVPA 239
DB 235 ADVSTGLSSGGSTLLSAMON--TVVWDVFPAPKE---RCVKTFQGNVH--FENL 288
QY 240 I--SPHNHNTATGSDQFVWIDPKNKLCQPHRYPTASLAFNSDGTLLAASS- 296
DB 289 LACSNSPDGSKIAAGSDRFFVYMDTSSRLYKLPGLHGSINVFAPHDPTIIASSD 348
QY 297 ---YKWDG 303
DB 349 XRLVMEFSE 358

```


TELEPHONE: (415) 324-0880
 TELEFAX: (415) 24-0960
 INFORMATION FOR SEQ ID NO. 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 TOPOLOG: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21
 US-08-190-802A-38

Query Match 9.68; Score 169.5; DB 1; Length 340;
 Best Local Similarity 24.88; Pred. No. 4.1e-10;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;
 QY 26 TSQFLVSS-----MD--TSRLVDPVANS--MLKLTQHTGAVLDCAFVDPYTHMS 72
 DB 65 TDSRLVSSAQCKLINSYTNKHAIFLSSHWMTCAFSQNVAC-----114
 QY 73 GOLDHOLKMDLTDQNI-----VQTHDAPICVEYCEPVNMYTSGNDQTVKMDPRT 127
 DB 115 GUNDKYCTMYLKTREGVNVRSLAGHTOYLSCCRFLAD--NOIVTSQDTCLMDIET 173
 QY 128 PNAGTGS--QPEKVTYLSVGD--RLVY--GTAGRRVYVMDLNRGVYQOR---RSSLKQY 181
 DB 174 GQQTTFYTGCDWSLAPDTLFCVGSACDASAKMDVRE--GNCRTQTFGHESI---229
 QY 182 TFCIRAPFKMGVGLSETEGRVAVYLDPSPEVQKVKAFKRLKENNIEQVYPAIS 241
 DB 230 -NALCFPGNAPATGSDDDATCRPLFLRAQDELMTYSH-----DN1--ICGITSVS 277
 QY 242 FHNHMTFATGSGDFWINDFENKRLCOHRVPTSIASLAFNSQDTLATAS 295
 DB 278 FSKSGRLLAGYDDFNCNFWMDALADKADRGVAGHDNRVSCLVGVDMDAVATGS 331

RESULT 10
 US-08-477-346-38
 Sequence 38, Application US/0847346
 Patent No. 6342369
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 51A
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY: MURASHIGE, KATE H.
 REFERENCE/DOCKET NUMBER: 29,959
 TELECOMMUNICATION INFORMATION:
 TELEFAX: (202) 887-0713
 INFORMATION FOR SEQ ID NO: 38:

TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO. 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 TOPOLOG: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21
 US-08-477-346-38

Query Match 9.68; Score 169.5; DB 3; Length 340;
 Best Local Similarity 24.88; Pred. No. 4.1e-10;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;
 QY 26 TSQFLVSS-----MD--TSRLVDPVANS--MLKLTQHTGAVLDCAFVDPYTHMS 72
 DB 65 TDSRLVSSAQCKLINSYTNKHAIFLSSHWMTCAFSQNVAC-----114
 QY 73 GOLDHOLKMDLTDQNI-----VQTHDAPICVEYCEPVNMYTSGNDQTVKMDPRT 127
 DB 115 GUNDKYCTMYLKTREGVNVRSLAGHTOYLSCCRFLAD--NOIVTSQDTCLMDIET 173
 QY 128 PNAGTGS--QPEKVTYLSVGD--RLVY--GTAGRRVYVMDLNRGVYQOR---RSSLKQY 181
 DB 174 GQQTTFYTGCDWSLAPDTLFCVGSACDASAKMDVRE--GNCRTQTFGHESI---229
 QY 182 TFCIRAPFKMGVGLSETEGRVAVYLDPSPEVQKVKAFKRLKENNIEQVYPAIS 241
 DB 230 -NALCFPGNAPATGSDDDATCRPLFLRAQDELMTYSH-----DN1--ICGITSVS 277
 QY 242 FHNHMTFATGSGDFWINDFENKRLCOHRVPTSIASLAFNSQDTLATAS 295
 DB 278 FSKSGRLLAGYDDFNCNFWMDALADKADRGVAGHDNRVSCLVGVDMDAVATGS 331

RESULT 11
 US-08-473-089-38
 Sequence 38, Application US/08473089
 Patent No. 6342369
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,089
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 29,959
 FILING DATE: 07-JUN-1995
 ATTORNEY: MURASHIGE, KATE H.
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEFAX: (202) 887-0713
 INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGIN: UNKNOWN
 INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21
 US-08-473-089-38

Query Match 9.64; Score 169.5; DB 4; Length 340;
 Best Local Similarity 24.81; Prefix 16.9;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;

Qy 26 TSQFLVSS-----MD--TSRLVDPVANS--MLKQYGTGAVLDCAFYDPFHAMS 72
 Db 65 TDSRLVSSQDGLLWSTYTNKVAIFLSSWMTCAVAFSGVAC-----114
 Qy 73 GLDHLQHLKMDLNTDQEL-----VTHDAPICVCEVAVVWVTGSMDQVTKLMDPT 127
 Db 115 GELDYSICYNLKTGKVVNSRELAGHTVLSCCFLDD-NQVTSGGDTTCALMDLET 173
 Qy 128 PCNAGTFS-OPERKVTYLSVSGD-RLIV-GTAGRVLMDLRNGYVQOR--RESSLKQY 181
 Db 174 GQQTTFTHGTGWNLSLADYVLPVSGCDASAKMDVRE-GMCRPTFGHESDI---229
 Qy 182 TCRIFAPNKGQVLSLSEGRVAVELDPSPQVKKKYAFKCHRLKENNIEQIYPVAIS 241
 Db 230 -NALCPFNAGNATGSDGTCFLPDLAAGELMTSH-----DNI--LCGITSVS 277
 Qy 242 FNIHTFNTATGSGDFWVDPNKKRLCOFHRYFTSLASLSDGTTLATAS 295
 Db 278 FSSGKLLAGTDGPNCAVMDALADKAGVLAGHDNRVSCVGLVTDGNAVATGS 331

RESULT 12-072A-38
 US Patent No. 6423684
 Sequence 38; Application US/08487072A
 Patent No. 6423684

GENERAL INFORMATION:
 APPLICANT: Nochi, Rosen, Darla
 INVENTOR: Nochi, Rosen, Darla
 TITLE OF INVENTION: MD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 ADDRESS: 265
 CO-ADDRESSES: Morrison & Forrester
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 TRANSMISSION TYPE: COPY disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: 08/487,072A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 REGISTRATION NUMBER: 2959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 897-1500
 TELEFAX: (202) 897-1500
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid

Query Match 9.64; Score 169.5; DB 4; Length 340;
 Best Local Similarity 24.81; Prefix 16.9;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;

Qy 26 TSQFLVSS-----MD--TSRLVDPVANS--MLKQYGTGAVLDCAFYDPFHAMS 72
 Db 65 TDSRLVSSQDGLLWSTYTNKVAIFLSSWMTCAVAFSGVAC-----114
 Qy 73 GLDHLQHLKMDLNTDQEL-----VTHDAPICVCEVAVVWVTGSMDQVTKLMDPT 127
 Db 115 GELDYSICYNLKTGKVVNSRELAGHTVLSCCFLDD-NQVTSGGDTTCALMDLET 173
 Qy 128 PCNAGTFS-OPERKVTYLSVSGD-RLIV-GTAGRVLMDLRNGYVQOR--RESSLKQY 181
 Db 174 GQQTTFTHGTGWNLSLADYVLPVSGCDASAKMDVRE-GMCRPTFGHESDI---229
 Qy 182 TCRIFAPNKGQVLSLSEGRVAVELDPSPQVKKKYAFKCHRLKENNIEQIYPVAIS 241

TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGIN: UNKNOWN
 INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21
 US-08-487-072A-38

Query Match 9.64; Score 169.5; DB 4; Length 340;
 Best Local Similarity 24.81; Prefix 16.9;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;

Qy 26 TSQFLVSS-----MD--TSRLVDPVANS--MLKQYGTGAVLDCAFYDPFHAMS 72
 Db 65 TDSRLVSSQDGLLWSTYTNKVAIFLSSWMTCAVAFSGVAC-----114
 Qy 73 GLDHLQHLKMDLNTDQEL-----VTHDAPICVCEVAVVWVTGSMDQVTKLMDPT 127
 Db 115 GELDYSICYNLKTGKVVNSRELAGHTVLSCCFLDD-NQVTSGGDTTCALMDLET 173
 Qy 128 PCNAGTFS-OPERKVTYLSVSGD-RLIV-GTAGRVLMDLRNGYVQOR--RESSLKQY 181
 Db 174 GQQTTFTHGTGWNLSLADYVLPVSGCDASAKMDVRE-GMCRPTFGHESDI---229
 Qy 182 TCRIFAPNKGQVLSLSEGRVAVELDPSPQVKKKYAFKCHRLKENNIEQIYPVAIS 241
 Db 230 -NALCPFNAGNATGSDGTCFLPDLAAGELMTSH-----DNI--LCGITSVS 277
 Qy 242 FNIHTFNTATGSGDFWVDPNKKRLCOFHRYFTSLASLSDGTTLATAS 295
 Db 278 FSSGKLLAGTDGPNCAVMDALADKAGVLAGHDNRVSCVGLVTDGNAVATGS 331

RESULT 13-099-1
 US Patent No. 6555522
 Sequence 13; Application US/09245039
 Patent No. 6555522

GENERAL INFORMATION:
 APPLICANT: Iyengar, Srinivas R.V.
 INVENTOR: Iyengar, Srinivas R.V.
 TITLE OF INVENTION: INTERACTING PROTEINS AND USES THEREOF
 NUMBER OF SEQUENCES: 2
 ADDRESS: 6923-074
 CO-ADDRESSES: Morrison & Forrester
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 TRANSMISSION TYPE: COPY disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: 09/09,245,039
 FILING DATE: 1999-02-05
 CLASSIFICATION: 2.0
 ATTORNEY/AGENT INFORMATION:
 REGISTRATION NUMBER: 2959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 897-1500
 TELEFAX: (202) 897-1500
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid

Query Match 9.64; Score 169.5; DB 4; Length 340;
 Best Local Similarity 24.81; Prefix 16.9;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;

Qy 26 TSQFLVSS-----MD--TSRLVDPVANS--MLKQYGTGAVLDCAFYDPFHAMS 72
 Db 65 TDSRLVSSQDGLLWSTYTNKVAIFLSSWMTCAVAFSGVAC-----114
 Qy 73 GLDHLQHLKMDLNTDQEL-----VTHDAPICVCEVAVVWVTGSMDQVTKLMDPT 127
 Db 115 GELDYSICYNLKTGKVVNSRELAGHTVLSCCFLDD-NQVTSGGDTTCALMDLET 173
 Qy 128 PCNAGTFS-OPERKVTYLSVSGD-RLIV-GTAGRVLMDLRNGYVQOR--RESSLKQY 181
 Db 174 GQQTTFTHGTGWNLSLADYVLPVSGCDASAKMDVRE-GMCRPTFGHESDI---229
 Qy 182 TCRIFAPNKGQVLSLSEGRVAVELDPSPQVKKKYAFKCHRLKENNIEQIYPVAIS 241


```

Db 496 QSPYGVYVSGHQRVARLWATHDYOPRI FAGHLADVKTFRFPNSNVATGSADTV 555
Oy 121 KUMDRTFCNAGTFSQPEKYVTLVSQDRLIVGTAGRVLWDLRMGYVQORRESSLKY 180
Db 556 RMD-----VLGN----- 564
Oy 181 OTRCIRAFPMKQVLSISIEGRVAVVYLDSPVQKKYAFKCHLKENNIEQIYPWNAI 240
Db 565 ---CVNFTGKG-----PNSL 579
Oy 241 SPNNHNTFATGSDGFVNIWDPFNKRLCOPHRYPTSIASLPSNDGTLTASWYE 300
Db 580 TSPNGRELKATDRLVNDIGHGLAVGELKGTDTYVCSATSPSGLIAGSS----- 634
Oy 301 MDTSEHPEGIFIROVDAET 321
Db 635 MDTVRLNDA--IAPEDLET 653

```

Search completed: November 12, 2003, 20:11:20
 Job time : 25 secs

OmGene version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:08:53 ; Search time 30 seconds
(without alignments)
1877,792 Million cell updates/sec

Title: US-10-084-700-2

Perfect score: 1772
Sequence: 1 MTCNSDFKLNQPPEDGISV.....DGIIFROVTDATKXKSPECT 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : Published Applications Aa.*

- 1: /cgn2_6/prodata/1/pubpaas/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaas/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaas/US06_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaas/US09_PUB.pep.*
- 5: /cgn2_6/prodata/1/pubpaas/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaas/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaas/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaas/US03_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaas/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubpaas/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaas/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubpaas/US09_PUBCOMB.pep.*
- 13: /cgn2_6/prodata/1/pubpaas/US09_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubpaas/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubpaas/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubpaas/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/1/pubpaas/US10_PUBCOMB.pep.*
- 18: /cgn2_6/prodata/1/pubpaas/US0_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a
score at least as high as the observed score. The result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	1772	100.0	328	10	US-09-095-881-5		Sequence 4, Appl
2	1772	100.0	330	9	US-09-095-881-5		Sequence 4, Appl
3	1772	100.0	332	9	US-09-095-881-5		Sequence 4, Appl
4	1747	98.6	330	14	US-10-084-700-27		Sequence 27, Appl
5	1743	98.4	326	14	US-10-084-700-29		Sequence 29, Appl
6	1743	98.4	326	14	US-10-084-700-30		Sequence 30, Appl
7	511.5	28.9	358	14	US-10-084-700-31		Sequence 31, Appl
8	509	28.7	358	14	US-10-084-700-30		Sequence 30, Appl
9	509	28.7	413	9	US-09-095-302-712		Sequence 712, PPT
10	509	28.7	413	9	US-09-095-302-712		Sequence 712, PPT
11	258	11.1	100	9	US-10-084-700-28		Sequence 28, Appl
12	197.5	11.1	100	9	US-09-095-299-1045		Sequence 1045, Appl
13	197.5	11.1	100	10	US-09-095-299-1045		Sequence 1045, Appl
14	197.5	11.1	472	14	US-10-077-111-1		Sequence 11, Appl
15	176.5	10.0	375	14	US-10-1119-352-1		Sequence 1, Appl

```

16 176 9.9 1356 14 US-10-077-111-10 Sequence 10, Appl
17 175.5 9.9 486 15 US-10-274-525-3 Sequence 3, Appl
18 169.5 9.6 340 10 US-09-324-2488-2 Sequence 2, Appl
19 169.5 9.6 340 10 US-09-324-2488-2 Sequence 2, Appl
20 166.5 9.4 281 12 US-10-245-618-20 Sequence 20, Appl
21 166.5 9.4 281 12 US-10-245-618-20 Sequence 20, Appl
22 166.5 9.4 540 10 US-09-213-888-7 Sequence 7, Appl
23 166.5 9.4 540 10 US-09-213-888-10 Sequence 10, Appl
24 166.5 9.4 540 10 US-09-328-877A-10 Sequence 10, Appl
25 166.5 9.4 540 12 US-10-245-618-14 Sequence 14, Appl
26 166.5 9.4 545 10 US-09-213-888-6 Sequence 6, Appl
27 166.5 9.4 553 10 US-09-328-877A-6 Sequence 6, Appl
28 166.5 9.4 553 10 US-09-328-877A-5 Sequence 5, Appl
29 166.5 9.4 553 10 US-09-328-877A-5 Sequence 5, Appl
30 166.5 9.4 553 12 US-10-245-618-8 Sequence 8, Appl
31 166.5 9.4 559 10 US-09-213-888-9 Sequence 9, Appl
32 166.5 9.4 559 10 US-09-213-888-9 Sequence 9, Appl
33 166.5 9.4 589 10 US-09-213-888-8 Sequence 8, Appl
34 166.5 9.4 589 10 US-09-328-877A-8 Sequence 8, Appl
35 166.5 9.4 626 10 US-10-245-618-18 Sequence 18, Appl
36 166.5 9.4 626 10 US-10-245-618-18 Sequence 18, Appl
37 166.5 9.4 627 12 US-09-328-877A-4 Sequence 4, Appl
38 166.5 9.4 627 12 US-09-328-877A-4 Sequence 4, Appl
39 166.5 9.4 627 10 US-09-213-888-21 Sequence 21, Appl
40 166.5 9.4 626 10 US-09-328-877A-21 Sequence 21, Appl
41 166.5 9.4 627 10 US-09-213-888-3 Sequence 3, Appl
42 166.5 9.4 627 12 US-10-245-618-5 Sequence 5, Appl
43 166.5 9.4 666 10 US-09-213-888-27 Sequence 27, Appl
44 166.5 9.4 666 10 US-09-328-877A-27 Sequence 27, Appl
45 166.5 9.4 669 10 US-09-213-888-25 Sequence 25, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-095-881-4
; Application US/0905881
; Sequence 4, Match 328, Length 328
; General Information:
; APPLICANT: Sealey, Todd
; TIME OF INVENTION: DETECTION OF LOSS OF THE WILD-TYPE
; FILLER: HUBBI GENES 10, 437
; CURRENT APPLICATION NUMBER: US/09/095, 881
; CURRENT FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 15
; SOURCE: Sealey, Todd
; SEQ ID NO 4
; LENGTH: 328
; TYPE: PPT
; ORGANISM: Homo sapien
US-09-095-881-4

```

```

Query Match
Best Local Similarity 100.0%; Score 1772; DB 10; Length 328;
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qv 1 MTCNSDFKLNQPPEDGISVFKFNTSQLLVSDTSVLDVDPVANSRLKQVHTGAVL 60
D 1 MTCNSDFKLNQPPEDGISVFKFNTSQLLVSDTSVLDVDPVANSRLKQVHTGAVL 60
Qv 61 DCAFPDPFHANGSGGLHQLKHLADQNIQVTHDADPFCVCEVNAVWVYVTSNQVTV 120
D 61 DCAFPDPFHANGSGGLHQLKHLADQNIQVTHDADPFCVCEVNAVWVYVTSNQVTV 120
Qv 121 KLVADPRTCAAGTSQPEKVTYLSVSGDLVTCVAGRVLWDLRMNGVYVQVRESSLAY 180
D 121 KLVADPRTCAAGTSQPEKVTYLSVSGDLVTCVAGRVLWDLRMNGVYVQVRESSLAY 180
Qv 181 QTCICFAPFNKQGVLSSTEGVAVLDPSPFQKKAFCRHKENNITQIYVPAWI 240
D 181 QTCICFAPFNKQGVLSSTEGVAVLDPSPFQKKAFCRHKENNITQIYVPAWI 240

```

Oy 121 SPHNHNTFATGSGPVNMDPNKKLCQPHRYPTSIASLAFSNDGTTLAASSYWE 300
 Db 241 SPHNHNTFATGSGPVNMDPNKKLCQPHRYPTSIASLAFSNDGTTLAASSYWE 300
 Oy 301 MDTPEHEDGIFRQVTDATKPSCT 328
 Db 301 MDTPEHEDGIFRQVTDATKPSCT 328
 RESULT 2 -700-2
 ; Seq ID No 1301
 ; Length: 332
 ; Sequence 27, Application US/10084700
 ; Publication No. US20020160403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Seelye, Todd
 ; TITLE OF INVENTION: HUBB3 GENE INVOLVED IN HUMAN CANCERS
 ; FILE REFERENCE: PP-01106.004/200130.438D1
 ; CURRENT APPLICATION NUMBER: US/10/084,700
 ; CURRENT FILING DATE: 2002-02-27
 ; PRIOR FILING DATE: 2000-09-05
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 328
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens
 US-10-084-700-2

Query Match
 Best Local Similarity 100.0%; Pred. No. 6,1e-180; Length 328;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MTGSNEFKLNPDEGIGSSVFSNPTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Db 1 MTGSNEFKLNPDEGIGSSVFSNPTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Oy 61 DCAFYDPTFHANGSGDLQKHMDLNTDOENLVTGHDAPRCVCEVAVVWVWVTSMDQTV 120
 Db 61 DCAFYDPTFHANGSGDLQKHMDLNTDOENLVTGHDAPRCVCEVAVVWVWVTSMDQTV 120
 Oy 121 KMDPRTFCNAGTFSQPKVYTVLSVSGDRLVGTAGRVLVMDLNNKYVQORRESSLY 180
 Db 121 KMDPRTFCNAGTFSQPKVYTVLSVSGDRLVGTAGRVLVMDLNNKYVQORRESSLY 180
 Oy 181 OTRCIRAFPNKQGVYLSIEGRVAVYLDPSPEVQKKYAFKCHLKENNIQIYPPVAI 240
 Db 181 OTRCIRAFPNKQGVYLSIEGRVAVYLDPSPEVQKKYAFKCHLKENNIQIYPPVAI 240
 Oy 241 SPHNHNTFATGSGPVNMDPNKKLCQPHRYPTSIASLAFSNDGTTLAASSYWE 300
 Db 241 SPHNHNTFATGSGPVNMDPNKKLCQPHRYPTSIASLAFSNDGTTLAASSYWE 300
 Oy 301 MDTPEHEDGIFRQVTDATKPSCT 328
 Db 301 MDTPEHEDGIFRQVTDATKPSCT 328

RESULT 3
 US-09-925-301-1301
 ; Seq ID No 1301
 ; Length: 332
 ; Sequence 27, Application US/09925301
 ; Patent No. US2002052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR FILING DATE: 2000-09-05
 ; PRIOR FILING DATE: 2000-09-05
 ; PRIOR FILING DATE: 2000-09-05
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1694
 ; SOFTWARE: Patensia Ver. 2.0

; Seq ID No 1301
 ; Length: 332
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-925-301-1301
 Query Match
 Best Local Similarity 100.0%; Pred. No. 6,1e-180; Length 332;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MTGSNEFKLNPDEGIGSSVFSNPTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Db 5 MTGSNEFKLNPDEGIGSSVFSNPTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 64
 Oy 61 DCAFYDPTFHANGSGDLQKHMDLNTDOENLVTGHDAPRCVCEVAVVWVWVTSMDQTV 120
 Db 61 DCAFYDPTFHANGSGDLQKHMDLNTDOENLVTGHDAPRCVCEVAVVWVWVTSMDQTV 124
 Oy 121 KMDPRTFCNAGTFSQPKVYTVLSVSGDRLVGTAGRVLVMDLNNKYVQORRESSLY 180
 Db 125 KMDPRTFCNAGTFSQPKVYTVLSVSGDRLVGTAGRVLVMDLNNKYVQORRESSLY 184
 Oy 181 OTRCIRAFPNKQGVYLSIEGRVAVYLDPSPEVQKKYAFKCHLKENNIQIYPPVAI 240
 Db 185 OTRCIRAFPNKQGVYLSIEGRVAVYLDPSPEVQKKYAFKCHLKENNIQIYPPVAI 244
 Oy 241 SPHNHNTFATGSGPVNMDPNKKLCQPHRYPTSIASLAFSNDGTTLAASSYWE 300
 Db 245 SPHNHNTFATGSGPVNMDPNKKLCQPHRYPTSIASLAFSNDGTTLAASSYWE 304
 Oy 301 MDTPEHEDGIFRQVTDATKPSCT 328
 Db 305 MDTPEHEDGIFRQVTDATKPSCT 332
 RESULT 4
 US-10-084-700-27
 ; Seq ID No 1301
 ; Length: 332
 ; Sequence 27, Application US/10084700
 ; Publication No. US20020160403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Seelye, Todd
 ; TITLE OF INVENTION: HUBB3 GENE INVOLVED IN HUMAN CANCERS
 ; FILE REFERENCE: PP-01106.004/200130.438D1
 ; CURRENT APPLICATION NUMBER: US/10/084,700
 ; CURRENT FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 330
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-084-700-27

Query Match
 Best Local Similarity 100.0%; Pred. No. 2,8e-177; Length 330;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MTGSNEFKLNPDEGIGSSVFSNPTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Db 1 MTGSNEFKLNPDEGIGSSVFSNPTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Oy 61 DCAFYDPTFHANGSGDLQKHMDLNTDOENLVTGHDAPRCVCEVAVVWVWVTSMDQTV 120
 Db 61 DCAFYDPTFHANGSGDLQKHMDLNTDOENLVTGHDAPRCVCEVAVVWVWVTSMDQTV 120
 Oy 121 KMDPRTFCNAGTFSQPKVYTVLSVSGDRLVGTAGRVLVMDLNNKYVQORRESSLY 180
 Db 121 KMDPRTFCNAGTFSQPKVYTVLSVSGDRLVGTAGRVLVMDLNNKYVQORRESSLY 190
 Oy 181 OTRCIRAFPNKQGVYLSIEGRVAVYLDPSPEVQKKYAFKCHLKENNIQIYPPVAI 240
 Db 181 OTRCIRAFPNKQGVYLSIEGRVAVYLDPSPEVQKKYAFKCHLKENNIQIYPPVAI 240

QY 241 SPHNHNTATGSGGVNIPDPNPKKQLCQPHRYPTSIASLAFSDNGTTLAASSWYE 300
 Db 241 SHHNNHTATGSGGVNIPDPNPKKQLCQPHRYPTSIASLAFSDNGTTLAASSWYE 300
 QY 301 MDTDEHPEDGIFIRQVTDATKPK 324
 Db 301 MDTDEHPEDGIFIRQVTDATKPK 324

RESULT 5
 ; Sequence 290-29
 ; Best Local Similarity 99.74; Pred. No. 7, 2e-17; Length 326;
 ; Publication No. US20020160403A1
 ; GENERAL INFORMATION:
 ; TITLE OF INVENTION: HUBB3 GENE INVOLVED IN HUMAN CANCERS
 ; FILE REFERENCE: PP-01406.004/200430.43BD1
 ; CURRENT APPLICATION NUMBER: US/10/084, 700
 ; CURRENT FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 29
 ; LENGTH: 326
 ; ORGANISM: Mus musculus
 US-10-084-700-29

Query Match 98.4%; Score 1919; DB 14; Length 326;
 Best Local Similarity 99.74; Pred. No. 7, 2e-17;
 Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTGSEFLKQPPEDGIESVFSPNTSQFLVSSMDTSVRLVDPANSLKYQHTGAVL 60
 Db 1 MTGSEFLKQPPEDGIESVFSPNTSQFLVSSMDTSVRLVDPANSLKYQHTGAVL 60

QY 61 DCAITDPTANSSGGLDQLQDNDLQDNLVGTADPACRVCYCFVNVWVTSQVQV 120
 Db 61 DCAITDPTANSSGGLDQLQDNDLQDNLVGTADPACRVCYCFVNVWVTSQVQV 120

QY 121 KMPDPTPCNAGTFSQPKVTLSSGDRVLVGTAGRVLWDLNNGVQVQRESLSKY 180
 Db 121 KMPDPTPCNAGTFSQPKVTLSSGDRVLVGTAGRVLWDLNNGVQVQRESLSKY 180

QY 181 QTRCIRAFPNQGVPLASIEGRVAEYLDPSVEVQKKYAFKCHKENI E01YPPVAI 240
 Db 181 QTRCIRAFPNQGVPLASIEGRVAEYLDPSVEVQKKYAFKCHKENI E01YPPVAI 240

QY 241 SPHNHNTATGSGGVNIPDPNPKKQLCQPHRYPTSIASLAFSDNGTTLAASSWYE 300
 Db 241 SPHNHNTATGSGGVNIPDPNPKKQLCQPHRYPTSIASLAFSDNGTTLAASSWYE 300

QY 301 MDTDEHPEDGIFIRQVTDATKPKS 325
 Db 301 MDTDEHPEDGIFIRQVTDATKPKS 325

RESULT 6
 ; Sequence 290-29
 ; Best Local Similarity 99.74; Pred. No. 7, 2e-17; Length 326;
 ; Publication No. US20020160403A1
 ; GENERAL INFORMATION:
 ; TITLE OF INVENTION: HUBB3 GENE INVOLVED IN HUMAN CANCERS
 ; FILE REFERENCE: PP-01406.004/200430.43BD1
 ; CURRENT APPLICATION NUMBER: US/10/084, 700
 ; CURRENT FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 32
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-084-700-32

Query Match 30.4%; Score 538; DB 14; Length 365;
 Local Similarity 95.71; Pred. No. 1, 7e-40;
 Matches 127; Conservative 59; Mismatches 135; Indels 26; Gaps 10;

QY 4 SNREFLKQPPEDGIESVFSPNTSQFLVSSMDTSVRLVDPANSLKYQHTGAVLDC 62
 Db 26 AUDTVLNSPDESIDTATSPQDVFSSSPKQVAINQVQPGQAGHSSPFLC 85

QY 63 AFY--DPTANSSGGLDQLQDNDLQDNLVGTADPACRVCY--CEPVNV--WVTSQM 116
 Db 63 TNSGDTKVASGCGNMLALQDNLVGTADPACRVCY--CEPVNV--WVTSQM 145

QY 117 DQTVLWMDPTPCNAGTFSQPKVTLSSGDRVLVGTAGRVLWDLNNGVQVQRES 176
 Db 146 DRTKYIMDRQDQPVSYVWVWVRVRYSMONGSLWATAEHIAINLANPTTIFAKTS 205

QY 177 SIAKYQTCIRAFPNQGVPLASIEGRVAEYLDPSVEVQKKY--YAFKCHKLK-----E 228
 Db 206 FUKMOTPCVACTMEWQATAGTASVGRCSINVIDG--MKKSGFSFKCHQTNPNRQPS 263

QY 229 NITQIYVPAISPHNHTATGSGGVNIPDPNPKKQLCQPHRYPTSIASL--AFS 285
 Db 264 NQSLVYPNFIATPHPLGTFTAGDGTNFMDKQERL--KGYPTLQASIPVCSFN 320

QY 286 NUGTTLAASSWY--ENDQTHPEGDIFIRQVTDATKPK 324
 Db 321 RNSGVFAVSLVQVHMQGNQPNPDYVPIRLIATHTDDEVEK 362

RESULT 7
 ; Sequence 31; Application US/10084700;
 ; Publication No. US20020160403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Seeley, Todd
 ; TITLE OF INVENTION: HUBB3 GENE INVOLVED IN HUMAN CANCERS
 ; FILE REFERENCE: PP-01406.004/200430.43BD1
 ; CURRENT APPLICATION NUMBER: US/10/084, 700
 ; CURRENT FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 31
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Schizosaccharomyces pombe
 US-10-084-700-31

Query Match 28.9%; Score 511.5; DB 14; Length 352;
 Best Local Similarity 34.8%; Pred. No. 1, 4e-45;
 Matches 115; Conservative 58; Mismatches 142; Indels 15; Gaps 7;

QY 6 EFLKQPPEDGIESVFSPNTSQFLVSSMDTSVRLVDPANSL--FLAYQHTGAVLDC 62
 Db 22 DVEVAQPPEDSISDLAFSQ--AEYLAASNSKSVRIEYQVQTSIGALYEHQGVLSY 80

QY 63 AF--YDPTANSSGGLDQLQDNDLQDNLVGTADPACRVCYCFVNV--WVTSQMD 118
 Db 81 WNRQGTQVASSGVSKSKYFDIQGVQAAHDAVCFVFEAKGTPTLATGSKDK 140

QY 119 TVKLWMDPTPCNAGTFSQPKVTLSSGDRVLVGTAGRVLWDLNNGVQVQRES 176
 Db 141 TLKWLQDRTSLATVSLFERYVAMCVHPLATVATERNICVNLSPBTKFLANSEL 200

QY 179 KYQTCIRAFPNQGVPLASIEGRVAEYLDPSVEVQKKYAFKCHKENI E01YPPVAI 238
 Db 201 AFQYISLACPIKDGATGAEVKATONIDENSSQ--NFSFCHRNQAGNSADVSN 258

QY 229 ALEFPHNHTATGSGGVNIPDPNPKKQLCQPHRYPTSIASLAFSDNGTTLAASSY-- 297
 Db 259 STAFQVGTSTGATGSGGVNIPDPNPKKQLCQPHRYPTSIASLAFSDNGTTLAASSY-- 318

QY 298 --WYENDQTHPEGDIFIRQVTDATKPK 324

1 NAME/KEY: SITE
 2 LOCATION: (12)
 3 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 4 NAME/KEY: SITE
 5 LOCATION: (16)
 6 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 7 NAME/KEY: SITE
 8 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 9 NAME/KEY: SITE
 10 LOCATION: (23)
 11 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 12 NAME/KEY: SITE
 13 LOCATION: (38)
 14 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 15 NAME/KEY: SITE
 16 LOCATION: (43)
 17 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 18 NAME/KEY: SITE
 19 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 20 NAME/KEY: SITE
 21 LOCATION: (49)
 22 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 23 NAME/KEY: SITE
 24 LOCATION: (53)
 25 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 26 NAME/KEY: SITE
 27 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 28 NAME/KEY: SITE
 29 LOCATION: (78)
 30 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 31 NAME/KEY: SITE
 32 LOCATION: (89)
 33 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 34 NAME/KEY: SITE
 35 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 36 NAME/KEY: SITE
 37 LOCATION: (99)
 38 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 39 US-09-726-239-1015

Query Match 11.1%; Score 197.5; DB 11; Length 100;
 Best Local Similarity 72.2%; Pred. No. 4.5e-13;
 Matches 3; Conservative 2; Mismatches 12; Indels 1; Gaps 1;
 Qy 44 VPANSMRLKQHTGAVLDCAQYTPHANSGLDQKRGHLDLTDQENLVGTHDA 97
 Db 24 VPANSMRLKQHTQAVLDCAQYTPHANSGLDQKRGHLDLTDQENLVGTHDA 76

RESULT 14-111-11
 Sequence 11, Application US/10077111
 Publication No. US20020187492A1
 GENERAL INFORMATION:
 APPLICANT: Teleduc, C. Gordon
 APPLICANT: Killema, Jill
 TITLE OF INVENTION: TEA
 PRIORITY REFERENCE: 3033-4114052
 CURRENT FILING DATE: 2002-02-15
 PRIOR FILING DATE: 2001-05-29
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 11
 LENGTH: 742

1 TYPE: PRT
 2 ORGANISM: Thermomonospora curvata
 3 FEATURE:
 4 ORGANIZATION: amino acid sequence encoded by the rRNA gene
 5 US-10-077-111-11
 Query Match 11.1%; Score 197.5; DB 14; Length 742;
 Best Local Similarity 27.7%; Pred. No. 8.8e-12;
 Matches 70; Conservative 48; Mismatches 112; Indels 23; Gaps 10;
 Qy 15 DGIATSKYKESNTQSLFELSVYSLVDPANSMELKQ-HTGAUVDCAFY-DPHAMS 72
 Db 502 DWIVATWSPED-GALLASSDQAVLMDVADAEAEAEFGHTYVDLAFSPGWSVAS 560
 Qy 73 GGLDGLQALMDLTDQENLV-GTHDAPTRCVCEVQVMMVTSMDQTVKMDPTICNA 131
 Db 561 GGGGTAKMLMDVATQGTGAVHGVIVAFSPGWSVASGGDTIRMDVATCKER 620
 Qy 132 QGTGSP-EKYVTLTSVSD-RLVGTAGRVLDPLANNVGVQCRSSLSKYTRICRAPP 189
 Db 621 DVLQANRVNLASLSPGSMVCHGSDSTVHLMVDVAQSELR-----TFEGHDVRAVA 675
 Qy 190 -NKGVLISLSEIRGAVELVDSPEVQKQAFKCHLKENNIEQLVPPVALSPHNHT 248
 Db 676 FPGDGLASSDQDTRMDVAVQSE-----HTLSEGT---PVISVAFPEGTT 724
 Qy 249 FATGQDQPNVM 261
 Db 725 LASASEDQIRIM 737

RESULT 15
 US-10-119-932-1
 Sequence 1, Application US/10119932
 Publication No. US20020186488A1
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Lel, Preeti
 APPLICANT: Cooper, Neil C.
 APPLICANT: Patterson, Chandra
 TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
 NUMBER OF SEQUENCES: 6
 PRIORITY REFERENCE: 3033-4114052
 ADDRESSER: Incyte Genomics, Inc.
 STREET: 3160 Porter Drive
 CITY: Palo Alto
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 OPERATING SYSTEM: DOS
 COMPUTER FILE COMPATIBLE
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 3.0
 CURRENT APPLICATION DATA: US/10/119, 932
 FILING DATE: 09-Apr-2002
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 PRIORITY REFERENCE: US/09/590,540
 FILING DATE: 08-Jun-2000
 APPLICATION NUMBER: 09/063,743
 FILING DATE: April 21, 1998
 ATTORNEY/AGENT INFORMATION:
 NAME: <unknown>
 REGISTRATION NUMBER: 43,168
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-845-4255
 TELEFAX: 650-845-4166
 INFORMATION FOR SEQ ID NO: 1:

```

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 375 amino acids
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ ORGANISM: Homo sapiens
/ CLONE: 051913
/ SEQUENCE DESCRIPTION: SEQ ID NO. 1:
US-10-119-512-1
Query March 10.08, Score 176.5, DB 14, Length 375;
Best Local Similarity 25.88; Pred No. 5'se-10;
Matches 80; Conservative 51; Mismatches 145; Indels 37; Gaps 13;
QY 14 EDCISSVSESTGSOFLIUSGHTVLYNDPA--NSKLIYKTSYAVLDCAF-YDPFH 69
DB 66 EGEVYCKFHNGST-LASAGFRLILANYGDCDVTATL-GHSGAVELHYTDSM 123
QY 70 ANSGELHQKAKHGLATQ--EMLVTHDAPTRCYEYCEFWNVAWTSMDQTYKMDPT 127
DB 124 LFSASTDKTVAMSESTGERVELAGTSPVNSCYFARGPOLACTGSDGTCGLNDIRK 183
QY 128 PCNAGTFQPEKV--YTLASVSGRLIVTGRGRVLMNLNMGYVQORRESSLYQTR-- 183
DB 184 KAAIQTFQTVGLAVTFNDTSQIISGIDNDIKVNL-----RQKLTITRGH 234
QY 184 ----CITAFPNQGVYLSISGRVAVEYLDPSVEVOKKKAFAKCHRLKNNIEQIYPNA 239
DB 235 ADSVYGLSLSSESTLISNANDN-TVWVDVRFAPKE-:::|:|:| 288
QY 240 I--SFNIIHTFATGSGDFNWDPNPKRKLQCFHYPTSIASLAFSNDGTLAASS- 296
DB 289 LKCSNPSGSKLAKSDRFTVYMTTSERLLYLPGHAGSINPAPHDFPIISASD 346
QY 297 ---VWVEDD 303
DB 349 KRLNGEFS 358

```

Search completed: November 12, 2003, 20:12:03
 Job time : 31 secs